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[7,14], Michal Slyper[4,15], Isaac Wakiro[7,16], Judit Jane-Valbuena[4,15], Orr Ashenberg[4], Asaf Rotem[7,17], Raphael Bueno[11], Orit Rozenblatt-Rosen[4,15], Kathleen Pfaff[9], Scott Rodig[10], Aaron N. Hata[12,13], Aviv Regev[4,5,15], Bruce E. Johnson[7], & Alexander M. Tsankov[1,2,3,4,18*]

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

24 [12]Massachusetts General Hospital Cancer Center, Charlestown, MA, USA

25 [13]Department of Medicine, Harvard Medical School, Boston, MA, USA

26 [14]Present address: Ginkgo Bioworks, 27 Drydock Avenue, Boston, MA, USA

27 [15]Present address: Genentech, 1 DNA Way, South San Francisco, CA, USA

28 [16]Present address: Louis V. Gerstner Jr. Graduate School of Biomedical Sciences at Memorial

29 Sloan Kettering Cancer Center, New York, NY, USA

30 [17]Present address: Astrazeneca, 35 Gatehouse Dr., Waltham, MA, USA

31 [18]Lead contact

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

33

ABSTRACT

TP53 is the most frequently mutated gene across many cancers and is associated with shorter survival in lung adenocarcinoma (LUAD). To define how *TP53* mutations affect the LUAD tumor microenvironment (TME), we constructed a multi-omic cellular and spatial tumor atlas of 23 treatment-naïve human lung tumors. We found that *TP53*-mutant (*TP53*^{mut}) malignant cells lose alveolar identity and upregulate highly proliferative and entropic gene expression programs consistently across resectable LUAD patient tumors, genetically engineered mouse models, and cell lines harboring a wide spectrum of *TP53* mutations. We further identified a multicellular tumor niche composed of *SPP1*$^{+}$ macrophages and collagen-expressing fibroblasts that coincides with hypoxic, pro-metastatic expression programs in *TP53*^{mut} tumors. Spatially correlated angiostatic and immune checkpoint interactions, including *CD274-PDCD1* and *PVR-TIGIT*, are also enriched in *TP53*^{mut} LUAD tumors, which may influence response to checkpoint blockade therapy. Our methodology can be further applied to investigate mutation-specific TME changes in other cancers.

50 **INTRODUCTION**

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

61

62 Tumor protein p53 (*TP53*) is the most commonly mutated gene in NSCLC (~50%) as well as
63 across many other cancers[8]. *TP53* mutations are more common in smokers[9], and are associated
64 with tumor progression and metastasis, leading to shorter survival in patients with NSCLC[10,11,12].
65 Despite being the most extensively studied tumor suppressor, *TP53* has been challenging to target
66 therapeutically due to the diversity of mutations reported and its involvement in a wide range of
67 regulatory pathways[13-15]. *TP53* is known to have an impact on the tumor microenvironment (TME)
68 during tumor development in mouse models[16-18]. Mutant *TP53* is also associated with increased
69 expression of PD-L1 and improved response to anti-PD1 therapy in human lung adenocarcinoma
70 (LUAD)[19-24], motivating us to investigate TME differences in LUAD patients with *TP53*-mutant
71 ($TP53^{mut}$) *vs.* *TP53*-wild type ($TP53^{WT}$) tumors.

72

73 Single cell and spatial profiling of multiple treatment-naïve, resectable NSCLC patient tumors
74 could advance our fundamental understanding of *TP53*-mutant cancer cells and their relationship
75 to the TME and provide insights that may yield new therapeutic opportunities but, prior studies
76 have not yet addressed this question. Single-cell RNA-sequencing (scRNA-seq) has previously
77 been used to characterize malignant cells and the TME in resectable NSCLCs, focusing on gene
78 expression differences between the primary tumor and adjacent normal tissue[25-28], during tumor
79 progression[29,30], or across metastatic sites[31,32]. Studies have also leveraged TME characteristics to
80 stratify NSCLC patients into clinically-relevant subgroups that could predict response to
81 immunotherapy[22,33]. One study focused on characterizing the TME and intra-tumoral
82 heterogeneity in *EGFR*^{mut} LUAD[34], but these previous scRNA-seq studies did not determine the
83 mutational status of *TP53* in tumors and did not investigate its impact on all cellular compartments
84 in the TME.
85
86 Here, we generated the most comprehensive, multi-omic atlas of NSCLC to date, combining
87 matched whole-exome sequencing, scRNA-seq, spatial transcriptomics, and multiplex
88 immunofluorescence from *TP53*^{mut} and *TP53*^{WT} human tumors, to identify *TP53* mutation-
89 associated changes in malignant cells and the TME. Our integrated atlas allowed us to identify
90 spatially validated, cell-cell interactions that may contribute to shaping the tumor architecture of
91 NSCLC. In doing so, we uncovered a hypoxic expression niche in *TP53*^{mut} LUAD and identified
92 possible mechanisms by which malignant tumor cells may limit the growth of blood vessels in the
93 TME. We further found a T cell compartment with highly exhausted features in *TP53*^{mut} *vs.*
94 *TP53*^{WT} LUAD, which was accompanied by an enrichment in multiple immune checkpoint
95 interactions between malignant, T, and myeloid cells. Finally, using a systematic approach to

96 detect cellular landmarks in spatial transcriptomics data, we observed a potential pro-metastatic
97 tissue niche consisting of *SPP1*-expressing tumor associated macrophages (TAMs) and a collagen-
98 expressing population of cancer-associated fibroblasts (CAFs) linked to hypoxia in the *TP53*^{mut}
99 LUAD TME. Overall, our approach provides a framework for leveraging multi-omic technologies
100 to understand the impact of specific genomic alterations[35,36] on the cellular and spatial architecture
101 that can enable discovery of new biomarkers and more optimal methods for patient stratification
102 in lung cancer and other tumor types.
103
104 **RESULTS**
105
106 **A multi-omic atlas of NSCLC**
107 To characterize the molecular, cellular, and spatial heterogeneity in NSCLC, we performed whole-
108 exome sequencing (WES), scRNA-seq, spatial transcriptomics (ST; 10X Visium), and multiplex
109 immunofluorescence (mIF) on primary tumor resections from 23 NSCLC patients (**Figure 1A**), of
110 which 18 were diagnosed as LUAD (**Table S1**). Twenty-one of the 23 patients had a history of
111 cigarette smoking (Median 40, Range 12-100 pack years) and all patients were confirmed to be
112 treatment-naïve prior to tumor resection. We used WES analysis to determine the total tumor
113 mutational burden, copy-number alterations (CNAs), and mutational status of genes commonly
114 altered in NSCLC, including *TP53*, *EGFR*, *KRAS*, *STK11, KEAP1*, *RBM10*, and *PTPRD* in each
115 tumor. There was high concordance between predicted CNA profiles from tumor-matched WES
116 and those inferred from scRNA-seq data (**Figure 1B** and **Figure S1A**), many of which overlapped
117 with commonly deleted (*e.g.*, *CDKN2A*, *PTPRD*, *B2M*) and amplified (*e.g.*, *NKX2-1*, *TERT*,
118 *MCL1*, *EGFR*) genes in LUAD[8]. After filtering, we obtained a total of 167,193 high-quality cell

119 profiles and used CNA inference to distinguish between malignant and non-malignant epithelial
120 cells and then integrated[37], clustered, and annotated them into 11 broad cell classes based on
121 coherent expression of known marker genes (**Figure 1C**,**D**; **Figure S1B**,**C**; **Table S2**).
122 Distinguishing tumors by their *TP53* mutation status, the different *TP53* mutations had comparable
123 effects in downregulating the average expression of *TP53* target genes in malignant cells relative
124 to *TP53*^{WT} LUAD tumors (**Figure 1E**). The remaining five tumors consisted of squamous,
125 mucinous, and colloid histology, and were not used for comparisons between *TP53*^{mut} and *TP53*^{WT}
126 tumors because *TP53* mutations have been found to have histology-dependent effects on patient
127 outcome[38].
128
129 ***TP53*^{mut} LUAD is associated with a distinct TME composition**
130 Cell subsets in the TME of *TP53*^{mut} *vs*. *TP53*^{WT} LUAD tumors had distinguishable cell-intrinsic
131 expression profiles (**Figure 1D**, **right**) and cell compositions (**Figure 1F** and **Figure S1D**). In
132 particular, there was a pronounced decrease in the proportion of endothelial cells and pericytes in
133 *TP53*^{mut} *vs*. *TP53*^{WT} LUAD tumors (**Figure 1F**, **left**), which we confirmed by scoring the
134 corresponding cell type signatures in 510 bulk RNA-seq profiles of primary LUAD tumors from
135 The Cancer Genome Atlas (TCGA) (**Figure 1F**, **right**). Endothelial cell abnormalities have been
136 linked to tumor cell progression and metastasis[39], and pericyte depletion has been linked to
137 hypoxia-associated epithelial to mesenchymal transition (EMT) in breast cancer[40].
138
139 These compositional differences between in *TP53*^{mut} and *TP53*^{WT} LUAD were also supported by
140 spatial transcriptomic profiling of 14 tissue sections from 6 tumors in our cohort (2 with *TP53*^{mut}
141 and 4 with *TP53*^{WT} LUAD). Spatial data from 2 additional tumors (6 slides) were excluded from

142 downstream analysis due to either having an non-LUAD histology or an insufficient number of
143 malignant cells (**Figure 1G**,**H** and **Table S1**). We used Tangram[41], a deep learning framework, to
144 map tumor-matched scRNA-seq profiles to the corresponding ST data and generate a probabilistic
145 measure of the cellular composition for each spatial measurement (**Figure 1G**). Spatial cell
146 composition inference was highly concordant with independent, manual annotations of the
147 corresponding hematoxylin and eosin (H&E) stains by a pulmonary pathologist (N.R.M.; **Figure**
148 **1G** and **Figure S1E-G**). Consistent with our scRNA-seq, there was also a decrease in the
149 proportion of endothelial cells and pericytes across ST spots in *TP53*^{mut} *vs*. *TP53*^{WT} tumor samples
150 (**Figure 1H**).
151
152 **Dichotomous activity of malignant cell programs in LUAD associates with survival**
153 While malignant cells from one tumor can occupy a gradient of expression states rather than
154 discrete cell subsets[42,43], different cancer expression programs and their usage across cells and
155 tumors have not been comprehensively defined in LUAD at single-cell resolution. To identify
156 shared cell states across tumors despite predominant, patient-specific differences in malignant
157 cells[42,43], we used canonical correlation analysis (CCA)[44] to integrate and cluster malignant cells
158 into 18 subsets shared across tumors (**Figure 2A** and **Figure S2A**) and annotated cell subsets using
159 gene set enrichment analysis (**Figure S2B**), highlighting representative markers (**Figure 2B**).
160 These subsets were enriched for genes associated with hallmark cancer processes (*e.g.*, cell cycle,
161 hypoxia[45-47], glycolysis[48], partial EMT (pEMT)[49], interferon gamma (IFNG) and TNFA signaling)
162 and with lung epithelial cell identity (*e.g.*, alveolar type 2 (AT2-like), ciliated, secretory). Other
163 malignant cell subsets showed marked expression of antigen presentation (MHCII), stress
164 response[50], metallothionein, and senescence-related genes[51], which have not been previously

165 identified *de novo* using scRNA-seq in NSCLC primary tumors. For downstream analyses, we
166 defined LUAD malignant programs by the set of 10 most differentially expressed genes for each
167 malignant subset and scored each program's gene set across all malignant cells.
168
169 The programs clustered into two main groups after correlating their scores across all malignant
170 cells (**Figure 2C**) and tumor samples (**Figure S2C**): one consisting of cell cycle, hypoxia, pEMT,
171 and glycolysis programs, and another of alveolar-like, antigen-presentation, and stress response
172 programs. Malignant program scores in these two groups were strongly anticorrelated, suggesting
173 that they were regulated by two distinct and opposing pathways (**Figure 2C** and **Figure S2C**).
174 This is consistent with a previous LUAD study where cells from patient-derived xenografts
175 consisted of two distinct intra-tumoral subgroups, distinguished by a gene module consisting
176 primarily of cell cycle genes[52]. High expression of several malignant programs (*e.g.*, pEMT,
177 hypoxia, cell-cycle, metallothionein, glycolysis) was associated with shorter survival across LUAD
178 patient tumors in TCGA (**Figure 2D** and **Figure S2D**). In contrast, high expression of the antigen-
179 presentation, MHCII program was linked to prolonged survival. Previously, various signatures
180 derived from bulk RNA profiling in TCGA have been shown to be predictive of outcome in
181 NSCLC[53,54] but our *de novo* approach allowed for the discovery of more granular programs based
182 on malignant cell expression only, unconfounded by other cells in the tumor microenvironment.
183
184 To identify cell-intrinsic differences between malignant cells from *TP53*^{mut} and *TP53*^{WT} LUAD,
185 we compared the mean activity score of different malignant programs between the two patient
186 groups and found higher expression in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD for several programs associated
187 with shorter survival (*e.g.*, cell cycle, glycolysis, pEMT), and a decreased expression of the AT2-

188 like program and *TP53* targets (**Figure 2E**, **left**), consistent with the observation that *TP53*^{mut}
189 tumors have shorter survival in LUAD TCGA (**Figure S2E**). There was not a statistically
190 significant difference in either the number of single-nucleotide alterations (SNAs) or CNAs
191 between *TP53*^{mut} and *TP53*^{WT} tumors in our cohort (**Figure S2F**). The malignant program
192 differences were corroborated in bulk RNA-seq data from 263 *TP53*^{mut} and 247 *TP53*^{WT} LUAD
193 primary tumors in TCGA (**Figure 2E**, **right**). These findings were also replicated when analyzing
194 a compendium of independent cohorts (**Table S3**), assembled from published and unpublished
195 scRNA-seq data from 22 additional treatment-naïve LUAD primary tumor samples with annotated
196 mutational status (7 *TP53*^{mut} and 15 *TP53*^{WT})[26,32,55] (**Figure S2G**). Notably, spatial co-expression
197 of cell cycle with glycolysis and of hypoxia with pEMT programs was higher in *TP53*^{mut} *vs*.
198 *TP53*^{WT} LUAD tumors (**Figure 2F**).
199
200 The decreased expression of the AT2-like program in *TP53*^{mut} LUAD was consistent with a loss
201 of AT2 cell identity. Scoring each malignant cell profile across tumors for normal lung epithelial
202 cell signatures[56], *TP53*^{WT} LUAD malignant cells scored highly for AT2 markers (**Figure 2G**),
203 whereas cells from tumors with squamous, neuroendocrine, or mucinous/colloid subtypes
204 expressed the expected corresponding basal, neuroendocrine, or goblet signatures[57]. Interestingly,
205 AT2 scores for *TP53*^{mut} LUAD malignant cells were significantly lower than *TP53*^{WT} LUAD but
206 were comparable to AT2 scores in the tumors from non-adenocarcinoma subtypes, suggesting a
207 loss of alveolar cell identity, which associates with tumor progression in LUAD[58].
208
209 **Consistent cancer-intrinsic changes observed across different *TP53* variant and co-mutation**
210 **classes**

211 To understand the potential joint effect of *EGFR* and *KRAS* co-mutations on the observed *TP53*^{mut}
212 LUAD cancer-intrinsic changes, we partitioned the TCGA cohort (n=506) into 6 subsets with
213 distinct mutational statuses for these three genes (**Figure S3A**, **left**). *TP53*^{mut} tumors with or
214 without co-mutations in *EGFR* or *KRAS* showed significant decreases in *TP53* target and AT2-
215 like scores, and a significant increase in CC.G2/M score compared to *TP53*^{WT}*KRAS*^{WT}*EGFR*^{WT}
216 tumors, whereas *TP53*^{WT} tumors with either *EGFR* or *KRAS* mutations did not have such
217 significant differences in these signatures compared to tumors that were *TP53*^{WT}*KRAS*^{WT}*EGFR*^{WT}.
218 Interestingly, tumors with *TP53* mutations alone or in combination with *KRAS* but not with *EGFR*
219 had significant increases in Glycolysis.hypoxia and pEMT scores compared to
220 *TP53*^{WT}*KRAS*^{WT}*EGFR*^{WT} tumors, suggesting that *EGFR* mutations counteracted the *TP53*-
221 induced effects on these malignant expression programs (**Figure S3A**), consistent with a recent
222 report[59]. Multiple regression analysis also showed consistent cancer-intrinsic changes associated
223 with *TP53* mutational status even when regressing out the effects of *EGFR*, *KRAS* (*G12C, G12D,*
224 *G12V,* as well as all *KRAS* mutations together)*, STK11*, *KEAP1*, *RBM10*, and *PTPRD* mutations
225 (**Figure 2H** and **Figure S3B**). Notably, there were no significant interaction effects between
226 mutational status of *TP53* and *KRAS* or *TP53* and *EGFR* after ANOVA testing of model outputs.
227
228 To investigate the potential impact of different *TP53* mutational variants on changes to the
229 expression of malignant gene programs, we partitioned the TCGA tumor samples based on
230 dominant negative (DNE), loss of function (LOF), and Impactful *TP53* mutational classes as
231 defined by previous functional studies overexpressing different *TP53*^{mut} variants in a *TP53*^{WT}
232 A549 LUAD cell line[60,61]. Consistent with our tumor scRNA-seq analysis, decreases in *TP53*
233 target, and AT2-like scores, and increases in CC.G2/M, Glycolysis.hypoxia, and pEMT scores

234 were associated with TCGA LUAD tumors with different sub-categories of *TP53* DNE and LOF
235 mutations (**Figure S3C**) or of *TP53* Impactful mutation sub-categories (**Figure S3D**). The
236 DNE,LOF and Impactful II mutations gave rise to the most pronounced differences, where
237 Glycolysis.hypoxia and pEMT programs are significantly increased only in DNE,LOF and
238 Impactful II variant categories in TCGA (**Figure S3C** and **Figure S3D**). Moreover, many of these
239 malignant cell intrinsic changes were also observed *in vitro* in A549 LUAD cells ($TP53^{WT}$)
240 overexpressing each of a broad spectrum of *TP53* mutational variants and profiled by Perturb-
241 seq[61] (**Figure S3E**). These trends held across all mutational categories, including all four *TP53*
242 variants also found in our patients' tumors and one variant also found in the scRNA-seq validation
243 compendium (**Figure S3E**). As in TCGA, the malignant expression program changes were most
244 pronounced for Impactful II mutations (**Figure S3E**) but were absent from A549 LUAD cells
245 overexpressing different *KRAS* mutations (**Figure S3F**). Because we see consistent changes in
246 malignant expression programs across a wide spectrum of *TP53* mutations, we reason that these
247 phenotypic changes are associated with a loss of $TP53^{WT}$ activity, rather than specific gain of
248 function that arises from any specific mutation. In support of our observations, heterogeneous
249 deletion events resulting in loss of *TP53* activity has recently been shown to result in homogenous
250 and deterministic patterns genome evolution[62].
251
252 Finally, we observed consistent changes in expression associated with *TP53* mutations at the
253 protein level. Analyzing bulk proteomic data from the Clinical Proteomic Tumor Analysis
254 Consortium (CPTAC) showed consistent decreases in *TP53* target and AT2-like signature scores,
255 and an increase in CC.G2M signature scores at the protein level (**Figure S3G**), where effects were
256 again most pronounced for DNE,LOF and Impactful II mutation categories (**Figure S3H**).

257

258 **Conserved changes in signaling entropy and malignant programs in *TP53*^{mut} human tumors**
259 **and mouse models**

260 We have previously reported loss of alveolar cell identity and reversion to a progenitor-like,
261 highly-plastic cell state (HPCS) during LUAD progression in mice with somatic activation of
262 oncogenic **K**RAS-G12D, which was greater in tumors from mice where the **p**53 was deleted (KP
263 model) than those with wild-type p53 (K model)[63]. Consistent with our findings in human tumors,
264 there was a significant decrease in AT2-like program score, and significant increases in CC.S,
265 Hypoxia, Glycolysis.hypoxia, pEMT program scores when comparing KP versus K mice
266 malignant cells (**Figure S4A**). Moreover, the expression of key hypoxia transcription factor *Hif1a*
267 was higher in KP *vs*. K malignant cells (**Figure S4A**, **right**). Interestingly, we did not observe
268 significant increases in the Hypoxia, Glycolysis.hypoxia, or pEMT programs between *TP53*^{mut} and
269 *TP53*^{WT} A549 LUAD cells *in vitro* (**Figure S3E**), suggesting that the tumor microenvironment is
270 important for shaping these malignant programs *in vivo*.

271

272 The consistent loss in alveolar type II cell identity in *TP53*^{mut} LUAD human tumors, cell lines, and
273 mouse models prompted us to explore the consequences of this loss on signaling entropy[64], a
274 representative measure of cellular plasticity. We found an increase in signaling entropy in *TP53*^{mut}
275 *vs. TP53*^{WT} LUAD malignant cells at the single-cell level in our cohort, which we confirmed in
276 bulk RNA-seq and bulk proteomic profiles across a larger cohort size (**Figure 2I**), as well as in
277 the KP *vs*. K mouse models[63] (**Figure 2J**). This increase in entropy was most significant in
278 DNE,LOF and Impactful II *TP53*^{mut} tumors in TCGA (**Figure S4B**) and was significant in *TP53*-

279 only mutants and *TP53-KRAS* co-mutants but not in *TP53-EGFR* co-mutants in TCGA (**Figure**
280 **S4C**), consistent with the trends observed in the malignant program activity changes.
281
282 Cell cycle and pEMT programs displayed the highest entropy amongst human malignant subsets
283 (**Figure S4D**) and the greatest similarity to the HPCS cell state described in the KP mouse model
284 (**Figure S4E**). Moreover, in A549 cells overexpressing mutant *TP53*, there was significantly
285 increased entropy in cells harboring Impactful I and Impactful II *TP53* mutations, including
286 variants found in our scRNA-seq and validation cohorts (**Figure 2K**). Conversely, there was no
287 significant change in entropy in the same cells overexpressing different classes of *KRAS* mutants
288 (**Figure S4F**), suggesting that the observed increase in entropy is directly linked to *TP53*
289 mutations. Taken together, malignant cells from *TP53*^{WT} LUAD tumors were enriched for antigen
290 presentation and AT2-like cell programs, while those from *TP53*^{mut} LUAD tumors were highly
291 entropic, plastic, and associated with programs predictive of poor outcome, including cell cycle,
292 glycolysis, hypoxia, and pEMT.
293
294 ***TP53*^{mut} LUAD is characterized by spatial interactions between malignant and stromal cells**
295 **known to inhibit vascularization and promote EMT**
296 We next focused on the decreased proportion of endothelial cells in *TP53*^{mut} LUAD (**Figure 1F**).
297 This association is also observed in *TP53* mutant tumors in TCGA LUAD (**Figure 1F**), with or
298 without co-mutations in *KRAS* and *EGFR* (**Figure S5A**, **left**), and is most significant in DNE,LOF
299 *TP53* mutants (**Figure S5A**, **right**).
300

301 To investigate the potential mechanisms for decreased vascularization, we annotated 9 endothelial
302 cell subsets based on the expression of established markers for endothelial cell subsets in the
303 normal lung (**Figure 3A**, **Figure S5B-D**)[65]. These included: aerocytes, a specialized capillary cell
304 responsible for gas exchange in the alveolus[66] (expressing *HPGD*, *EDNRB*, and *SOSTDC1*);
305 arterial cells (*GJA5*, *ENPP2*); capillary cells (*CA4*, *FCN3*); vascular endothelial cells (VEC),
306 including *COL4A1*$^{+}$ VECs (VEC.COL4A1), cycling (VEC.Cycling) and interferon-stimulated
307 (VEC.IFN) subsets; lymphatic cells; pulmonary venous cells; and systemic venous cells. The
308 proportion of aerocytes — and to a lesser degree arterial cells — out of all endothelial cells was
309 decreased in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD (**Figure 3B**), which was further supported by bulk RNA-
310 seq data deconvolved using scRNA-seq-derived markers (**Figure S5E**). This suggests a
311 corresponding decrease in gas exchange in *TP53*^{mut} LUAD and a resulting decrease in oxygen
312 diffusion to the area directly surrounding the alveoli at the tumor site, providing a selective
313 advantage for tumor cell survival[67].
314
315 Analyzing ligand-receptor pairs, several interactions between malignant and endothelial cells
316 known to inhibit endothelial cell growth and function were enriched in *TP53*^{mut} LUAD, including
317 *SEMA3A-NRP1* and *EPHB2-EFNB1*[68-70] (**Figure 3C** and **Figure S5F)**. Spatial transcriptomics
318 from matched tumor samples support these observations, showing a significantly higher spatial
319 correlation in the expression of *SEMA3A-NRP1* and *EPHB2-EFNB1* ligand-receptor pairs in
320 *TP53*^{mut} *vs*. *TP53*^{WT} LUAD sections (**Figure 3D** and **Figure S5G**). Taken together, we observe
321 spatially correlated expression of known inhibitory interactions enriched between malignant and
322 endothelial cells in *TP53*^{mut} LUAD, which may explain the accompanying decrease in these
323 subsets and vascularization.

324

325 The heterogeneity of mesenchymal cells and their potential role in faster tumor progression and
326 shorter patient survival has been documented in NSCLC[26,71]. Mesenchymal cells in our tumors
327 were partitioned into 13 subsets (**Figure 3E** and **Figure S6A-C**). Two subsets most closely
328 resembled normal lung myofibroblasts[56] (**Figure 3F**): CAF.COLs expressing high levels of
329 collagens (*COL1A2*, *COL3A1*, *COL1A1*), which share markers with recently described *LRRC15*$^{+}$
330 myofibroblasts[72], and myofibroblasts, expressing both fibroblast and smooth muscle cell markers.
331 We also annotated two pericyte-like subsets, one expressing canonical pericyte markers (*HIGD1B*,
332 *COX4I2*), and one expressing *COL4A2* and EMT-related genes (**Figure S6A**,**B**); airway and
333 vascular smooth muscle subsets, and alveolar (CAF.ADH1B) and adventitial (CAF.Adventitial)
334 fibroblasts, which correspond to recently described human lung alveolar and adventitial
335 fibroblasts[72]; and fibroblast subsets expressing higher levels of lipid metabolism (CAF.APOE),
336 ribosomal (CAF.Ribo), complement and chemotaxis (CAF.Complement), and interferon-
337 stimulated (CAF.ISGs) genes. Of note, we did not observe CAF subsets that resembled normal
338 lung fibromyocytes or lipofibroblasts (**Figure 3F**). The different mesenchymal subsets were
339 associated with different active regulons (**Figure 3G**), as inferred by pySCENIC[73], with high
340 regulon specificity (**Figure S6D**). As expected, cycling CAFs had an active *E2F1* regulon, a cell
341 cycle regulator; CAF.ISGs activated the *IRF7* and *STAT1* regulons, controlling interferon
342 responsive genes; and CAF.COLs activated a *TWIST1* regulon, previously linked to EMT in
343 cancer[74].

344

345 Based on the inferred composition of CAF subsets in spatial transcriptomics spots across sections,
346 pericytes were significantly spatially correlated with endothelial cells (**Figure S6E**, Pearson

347 correlation *P*-value < 10^{-15}); CAF.COLs with vascular smooth muscle and CAF.ADH1B,
348 CAF.Complement with CAF.Adventitial, pericytes with CAF.ISGs, and airway smooth muscle
349 cells with CAF.APOE and CAF.Ribo. Spatially, most mesenchymal subsets were negatively
350 correlated with malignant cells, indicating occlusion of fibroblasts from the tumor core (**Figure**
351 **S6E**). Comparing the relative proportions of mesenchymal cell subsets (out of all mesenchymal
352 cells) between *TP53*^{mut} *vs*. *TP53*^{WT} LUAD, pericytes and CAF.ADH1B were decreased in *TP53*^{mut}
353 LUAD (as observed for broad cell classes) (**Figure 3H**), also supported by deconvolved bulk
354 RNA-seq (**Figure 1F** and **Figure S6F**). Similar to endothelial scores, pericyte scores were
355 significantly decreased in *TP53* mutant tumors in TCGA LUAD, with or without co-mutations in
356 *KRAS* and *EGFR*, and this decrease was most significant in the DNE,LOF class of *TP53* mutations
357 in TCGA (**Figure S6G**).
358
359 *TGFB2-TGFBR2* had significantly higher spatial correlation in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD
360 (**Figure 3I**), and *TGFB2-TGFBR2* and *TGFBR2-TGFB3* interactions between mesenchymal and
361 malignant cells were enriched in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD (**Figure S6H**), although the
362 enrichment was not significant. *VEGFA* expression in myofibroblasts is higher in *TP53*^{mut} *vs*.
363 *TP53*^{WT} LUAD (**Figure S6I**), consistent with previous reports that myofibroblasts can regulate
364 angiogenesis in response to hypoxia[75]. Overall, changes in both stromal cell composition and
365 spatial organization in *TP53*^{mut} LUAD are accompanied by an enrichment of putative interactions
366 known to inhibit vasculature and induce epithelial-mesenchymal transition.
367
368 **Increased *SPP1* expression and immunomodulatory interactions in *TP53*^{mut} LUAD myeloid**
369 **cells**

370 Myeloid cells play an important role in tumor progression and may be a compelling therapeutic
371 target in cancer[76]. While NSCLC myeloid cell populations were generally characterized[22,27], the
372 impact of *TP53* mutations on myeloid cells, and their interactions with malignant, stromal, and
373 immune compartments in the TME has not been studied in detail in humans.
374
375 Myeloid cells partitioned into 14 subsets, largely consistent with previous NSCLC atlases[22]
376 (**Figure 4A,B** and **Figure S7A**), but with significant compositional differences between *TP53*^{mut}
377 and *TP53*^{WT} LUAD. These include a significant decrease in *TP53*^{mut} LUAD in the proportion of
378 TAM.FABP4 cells, which resemble tissue-resident alveolar macrophages (TAMs), and an increase
379 in *SPP1*-expressing TAMs (TAM.SPP1) (**Figure 4C**). In agreement, *SPP1* expression was
380 significantly increased both in TCGA bulk RNA-seq data and more prominently in myeloid
381 pseudobulk comparisons between *TP53*^{mut} and *TP53*^{WT} LUAD tumors (**Figure 4D**, **left**).
382 Expression of *CXCL9/10/11* was also increased in *TP53*^{mut} bulk TCGA LUAD tumors (**Figure**
383 **S7B**). The proportion of *CXCL9/10/11* expressing TAMs (TAM.CXCLs) was also higher in
384 *TP53*^{mut} LUAD tumors both in the scRNA-seq data and in deconvolution analysis of TCGA bulk
385 RNA-seq data (**Figure S7C**). Notably, increases in *SPP1* and *CXCL9/10/11* were most significant
386 in *TP53* mutant tumors without co-mutations in *EGFR* or *KRAS*, and in DNE,LOF and Impactful
387 II classes of *TP53* mutant tumors in TCGA (**Figure S7D,E**).
388
389 Myeloid cell subsets in *TP53*^{mut} and *TP53*^{WT} LUAD were also associated with distinct cell-
390 intrinsic gene programs, such that genes differentially expressed between myeloid subsets from
391 *TP53*^{mut} *vs*. *TP53*^{WT} tumors were enriched for EMT, hypoxia, and interferon response pathway,
392 especially in TAM.SPP1 cells (**Figure 4E** and **Figure S7F**). *SPP1* expression has been previously

393 associated with EMT, and with early lymph node metastasis in LUAD[77], as well as with shorter
394 survival in TCGA LUAD independent of *TP53* mutational status (**Figure S7G**). Notably, the genes
395 most correlated with *SPP1* expression across single myeloid cells were enriched for EMT, hypoxia
396 and glycolysis functions (**Figure 4D**, **right**) and included VEGF receptors *NRP1* and *FLT1*
397 (**Figure S7H**). In support of this observation, a recent compendium of 399 published scRNA-seq
398 studies has shown that *SPP1*$^{+}$ macrophages are consistently found in fibrotic lung diseases,
399 COVID-19 lung, and pancreatic ductal adenocarcinoma (PDAC), suggesting a role of these
400 fibrosis-associated *SPP1*$^{+}$ macrophages in lung injury and ECM remodeling[78]. In summary, we
401 find increased expression of *SPP1* in *TP53*^{mut} LUAD myeloid cells that is tightly linked with
402 induction of genes involved in EMT and hypoxia programs.
403
404 To identify ligand-receptor interactions that may contribute to a distinct myeloid compartment in
405 *TP53*^{mut} LUAD, we conducted differential ligand-receptor interaction analysis. Putative
406 interactions enriched in *TP53*^{mut} LUAD include *VEGFA-FLT1* and *NRP1-SEMA3A* (**Figure 4F**),
407 involved in recruitment of myeloid cells to tumor sites[79], and *TGFB2-TGFBR1* interactions
408 between malignant and myeloid cells, which are known to promote monocyte differentiation into
409 TAMs[80], and may help explain the relative increase in the monocyte-macrophage population in
410 *TP53*^{mut} LUAD. Additionally, *NRP1-VEGFA*, *NRP1-VEGFB*, *FLT1-VEGFB*, and *FLT1-PGF*
411 interactions between myeloid and mesenchymal cells, induced in *TP53*^{mut} LUAD, suggest that
412 mesenchymal cells may contribute to monocyte recruitment to the TME (**Figure 4F**)[79]. In addition,
413 *CXCL11-CXCR3* interactions between myeloid and B/T cells were enriched in *TP53*^{mut} LUAD
414 (**Figure 4F**) with *CXCR3* specifically expressed by NK, B, and T cells (**Figure 4G**) and *CXCL11*
415 induced in multiple myeloid subsets in *TP53*^{mut} LUAD (**Figure 4H**), most highly in TAM.CXCLs

416 (**Figure S7I**), which constitute a higher proportion of myeloid cells in *TP53*^{mut} LUAD. Thus,
417 myeloid cells may enhance recruitment of *CXCR3*-expressing T cells in *TP53*^{mut} LUAD. Finally,
418 there was a significant increase in putative *SPP1*-mediated interactions between myeloid and
419 endothelial cells or pericytes in *TP53*^{mut} LUAD (**Figure 4F**), also supported by increased spatial
420 correlation of *SPP1* and *CD44* expression in *TP53*^{mut} tumors (**Figure 4I**). Thus, *TP53*^{mut} LUAD
421 has increased expression of putative interactions involved in monocyte and lymphocyte
422 recruitment, TAM differentiation, hypoxia, EMT, and other downstream effector functions *via*
423 *SPP1, CXCL11*, and other TAM-associated immunomodulators.
424

425 **Enrichment of T cell exhaustion programs and putative immune checkpoint interactions in**
426 ***TP53*^{mut} LUAD**

427 Lymphoid cells in the scRNA-seq data partitioned into 13 subsets of T and NK cells (**Figure 5A,B**
428 and **Figure S8A**), follicular and marginal zone B cells, IgA and IgG plasma cells, plasmacytoid
429 DCs, and mast cells (**Figure S8B**). While there were no significant differences in B cell
430 composition, *TP53*^{mut} LUAD had a decreased proportion of *CD4*$^{+}$ tissue-resident memory T cells
431 (CD4.TRM) and an increase in pre-exhausted-like[81], *GZMK*-expressing *CD8*$^{+}$ (CD8.GZMK) and
432 exhausted-like (T.Exhausted) T cell populations out of all T and NK cells (**Figure 5C**), which was
433 supported by our scRNA-seq validation compendium (**Figure S8C**) and by deconvolved bulk
434 RNA-seq data from TCGA primary tumors (**Figure S8D**). Higher abundance of exhausted T cells
435 may be mediated by increased chemotaxis *via CXCL11-CXCR3* interactions (**Figure 4F-H**), where
436 *CXCR3* was upregulated in *TP53*^{mut} exhausted T cells (**Figure S8E**). We also observed a higher
437 proportion of TFH cells relative to all T and NK cells in *TP53*^{mut} *vs. TP53*^{WT} tumors in both our
438 discovery and validation scRNA-seq cohorts (**Figure S8C**), which was statistically significant

439 when combining all samples. Expression of *CXCL13*, a top predictor for response to checkpoint
440 inhibition therapy[82], was most highly expressed in exhausted-like and TFH subsets and
441 significantly increased in *TP53*^{mut} tumors (**Figure S8E**). Furthermore, expression of immune
442 checkpoint molecules *PDCD1*, *CTLA4*, *HAVCR2*, and *TIGIT* were significantly higher across
443 multiple T cell subsets in *TP53*^{mut} *vs.* *TP53*^{WT} tumors, while expression of *KLRB1*, a marker of
444 favorable prognosis in cancer[83], was downregulated (**Figure S8E**). Taken together, increased
445 expression of *CXCL13* and immune checkpoint molecules (*PDCD1*, *CTLA4*, *HAVCR2*, and
446 *TIGIT*) and proportions of TFH, *GZMK*-expressing *CD8*$^{+}$, and exhausted-like T cells all suggest a
447 heightened immunogenic potential in the *TP53*^{mut} LUAD TME that may contribute to a more
448 favorable ICI immunotherapy response[82,84-86].
449
450 Ligand-receptor analysis of putative immune checkpoint interactions[87] between T, myeloid, and
451 malignant cells showed an enrichment in *PDCD1-CD274* and *TIGIT-PVR* putative interactions
452 between T and malignant cells in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD (**Figure 5D**) and in *CTLA4-CD86*
453 and *HAVCR2-LGALS9* putative interactions between T and myeloid cells. Indeed, we found higher
454 expression of *CD274*, *CD86*, *PVR*, *PDCD1*, *CTLA4*, and *TIGIT* expression in *TP53*^{mut} bulk RNA-
455 seq LUAD tumor samples from TCGA (**Figure S8F**), higher expression of *PVR* and *CD274* in
456 *TP53*^{mut} bulk proteomic LUAD tumor samples from CPTAC (**Figure S8G**), and higher expression
457 of *PDCD1*, *CTLA4*, and *TIGIT* in multiple T cell subsets in our scRNA-seq data (**Figure S8E**,
458 **right**), where *TIGIT* was the most frequently expressed across T cells compared to *CTLA4* and
459 *PDCD1* (**Figure 5E**). *PVR* expression was also higher in malignant cells from *TP53*^{mut} *vs*. *TP53*^{WT}
460 tumor scRNA-seq data (**Figure 5F**, **left**) and highly correlated with the expression of the cell cycle
461 program in malignant cells (**Figure 5F**, **right**), suggesting that *PVR* may be activated during cancer

462 cell proliferation. Furthermore, there was significantly higher spatial colocalization of *TIGIT* and
463 *PVR* expression in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD (**Figure 5G**). In agreement, there was also a higher
464 proportion of *PVR*$^{+}$ malignant cells in *TP53*^{mut} *vs*. *TP53*^{WT} tumors (**Figure 5H**, **top**) and even more
465 prominently in cycling malignant cells (**Figure 5H**, **bottom**).
466
467 To investigate the *TP53*^{mut} enrichment in interactions between PD-L1 (*CD274*) and PD1 (*PDCD1*)
468 at a single-cell resolution, we applied multiplex immunofluorescence (mIF) to patient-matched
469 *TP53*^{mut} and *TP53*^{WT} formalin-fixed paraffin-embedded (FFPE) slides from our cohort (**Figure 5I**
470 and **Figure S8H**). There was a significant increase in log count density (cells/mm^2) of PD-
471 L1^{+}cytokeratin^{+} cells, as well as PD1^{+}CD8^{+} cells in *TP53*^{mut} *vs*. *TP53*^{WT} tumor samples (**Figure**
472 **5J**), and, consistent with the scRNA-seq cell-cell interaction inference (**Figure 5D**), an increase in
473 colocalization between PD-L1^{+}cytokeratin^{+} cells and PD1^{+}CD8^{+} cells in *TP53*^{mut} *vs*. *TP53*^{WT}
474 tumor samples, across multiple spatial analyses measuring direct contact, adjacency (as defined by
475 cells within the 6 nearest neighbors), and proximity (<50 μm) (**Figure 5K**). Taken together,
476 increased expression of *PVR*, *TIGIT*, *CD274*, and *PDCD1* was concurrently observed with
477 increased colocalization between the ligands and corresponding receptors in *TP53*^{mut} LUAD.
478
479 **Spatially-defined fibroblast-macrophage niche enriched in hypoxia and EMT programs in**
480 ***TP53*^{mut} LUAD**
481 To comprehensively map inter-compartment cellular colocalization using the spatial data, we
482 deconvolved the proportion of cell subsets across all spatial transcriptomics spots and calculated
483 the average correlation of cell type proportions across all LUAD sections and their difference
484 between *TP53*^{mut} and *TP53*^{WT} LUAD tumors (**Figure 6A**, **Figure S9A**). Interestingly, there was a

485 significantly lower colocalization (*i.e.,* more negative correlation) of endothelial cells with
486 mesenchymal cells and with myeloid cells in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD (**Figure 6B**), consistent
487 with the increase in hypoxia related gene expression in mesenchymal and myeloid subsets in
488 *TP53*^{mut} tumors. In addition, spatial colocalization of B, Plasma, and malignant cells was increased
489 in *TP53*^{mut} LUAD (**Figure S9B**). B cell infiltration has been associated with increased PD1/PD-
490 L1 expression, TMB, and prolonged survival in NSCLC[88] and likely contributes to the heightened
491 immunogenicity of *TP53*^{mut} LUAD tumors.
492
493 We further identified multicellular communities by an NMF-based method to derive 15 factors
494 with varying participation of different cell subsets (**Figure 6C**). Factor 7 (NMF7) contained
495 multiple myeloid and mesenchymal subsets with known significance in tumor progression,
496 including TAM.SPP1, CAF.COLs, and myofibroblasts, which were highly colocalized with each
497 other, as well as with hallmark hypoxia and EMT program expression (**Figure 6D**). Spots high in
498 NMF7 were low in malignant cell proportions and high in myeloid and mesenchymal cell
499 proportions (**Figure S9C**). NMF7 was also more positively correlated with hallmark hypoxia and
500 EMT program expression in the tumor periphery (spots with ≤50% malignant cells) and more
501 negatively correlated with endothelial cells across all spots in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD (**Figure**
502 **6E**). Because CAFs and macrophages are both suggested to induce EMT in tumors[89,90], NMF7
503 may represent a hypoxic fibroblast-macrophage niche that resides on the periphery of the tumor
504 core and may contribute to the early EMT phenotype in malignant cells in *TP53*^{mut} LUAD.
505
506 **DISCUSSION**
507

508 Targeted therapies for NSCLC are currently directed at oncogenic driver mutations and
509 chromosomal rearrangements (*e.g.*, *EGFR*, *KRAS-G12C, ALK*), yet there are currently no effective
510 therapies targeting *TP53* mutations, despite being common in tumors and extensively studied.
511 Restoration of *Trp53* expression has been shown to cause tumor regression in mice[91], leading to
512 efforts to treat *TP53*^{mut} tumors by restoring wild-type p53 activity[14]. We hypothesized that cellular
513 and spatial profiling of the TME in *TP53*^{mut} LUAD could provide a more detailed understanding
514 of how *TP53* mutations could contribute to shorter survival and suggest new therapeutic
515 opportunities. To this end, we characterized the LUAD TME of *TP53*^{mut} *vs*. *TP53*^{WT} tumors by
516 combining scRNA-seq of 167,193 cells across 23 NSCLC patients, 23 matched tumor and normal
517 WES, 20 spatial transcriptomics and 4 mIF tissue sections. Together, these represent the most
518 comprehensive multi-omic map of NSCLC to date and provide a valuable resource of cell
519 annotations, markers, and integrated data sets for the scientific community.
520
521 Our integrative analysis revealed a substantial remodeling of the TME associated with *TP53*
522 mutations (**Figure 6F**), including decreased presence of pericytes and endothelial aerocytes. The
523 mechanisms of TME remodeling are likely complex but our study uncovers multiple ones that may
524 play a role. These include putative ligand-receptor interactions between malignant and endothelial
525 cells (*SEMA3A-NRP1*, *EPHB2-EFNB1*) that were enriched in *TP53*^{mut} tumors and have known
526 roles in inhibiting vascularization, and which were significantly more colocalized in the *TP53*^{mut}
527 spatial data. *TP53*^{mut}-specific depletion of endothelial cells was accompanied by increased hypoxia
528 and pEMT program expression in malignant cells, which was altered across multiple functional
529 categories of *TP53* mutations in bulk transcriptomic and proteomic data, most prominently in
530 DNE,LOF and Impactful II *TP53* mutations. This effect was independent of presence or absence

531 of *KRAS* mutations but was partially reduced in tumors with *TP53-EGFR* co-mutations. Hypoxia
532 has been shown to induce EMT through *HIF-1α* as well as other mediators[92,93] and to positively
533 select for p53-deficient cells that have lost apoptotic potential in mice[94]. Thus, a hypoxic TME
534 may provide a survival advantage for highly plastic *TP53*^{mut} LUAD malignant cells, which have
535 lost alveolar identity and experience increased proliferative capacity and metastatic potential. This
536 model is supported by our analysis of data from genetically engineered K and KP mouse models,
537 and from A549 LUAD cell line (*TP53*^{WT}) overexpressing different *TP53* mutational variants.
538 Furthermore, *de novo* discovery of spatial multicellular communities revealed a highly hypoxic,
539 EMT-promoting spatial niche enriched for *SPP1*$^{+}$ TAMs, myofibroblasts, and collagen producing
540 CAFs. *SPP1* expression was highly upregulated in *TP53*^{mut} monocytes and TAMs and tightly
541 linked with regulation of hypoxia and EMT. Increased TGF-β signaling likely also helps shape an
542 EMT-supporting niche in *TP53*^{mut} tumor samples, where *TGFB2-TGFBR2* interactions involving
543 malignant cells were significantly enriched in spatial data. Previous studies linking mutant p53 to
544 metastatic potential have focused on malignant cell-intrinsic processes[11,12,95], as the effects of
545 mutant p53 on tumor-stromal and tumor-immune crosstalk have been described only recently in
546 model systems[96]. Our study is the first to profile the cellular and spatial context of *TP53* mutated
547 LUAD tumors to provide insight into how *TP53* mutations could remodel the TME to promote
548 tumor survival and metastasis in patients, leading to poorer outcomes.
549
550 Interestingly, one aspect of TME remodeling in *TP53*^{mut} LUAD is a higher tumor infiltration by B
551 cells and *CD8*$^{+}$ T cells. Increased lymphocyte recruitment may be caused in part by enriched
552 *CXCL11-CXCR3* ligand-receptor interactions between myeloid and B/T cells in *TP53*^{mut} tumors,
553 most notably exhausted T cells. We observed some overlap of the myeloid and T cell phenotypes

554 in *TP53*^{mut} LUAD, the anti-tumor immunity hub we reported previously in colorectal cancer[97], and
555 the stem immunity hub in NSCLC[98], but additionally report here B/Plasma cell tumor infiltration,
556 T cell exhaustion (supported by an independent scRNA-seq validation compendium), and the
557 enrichment of *SPP1*-expressing TAMs. Our findings of a potentially more immunogenic TME in
558 *TP53*^{mut} LUAD is consistent with previous studies reporting higher levels of anti-tumor immune
559 signatures[99] in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD using bulk RNA-seq data from TCGA and longer
560 survival in advanced *TP53*^{mut} *vs*. *TP53*^{WT} NSCLC patients receiving immune checkpoint
561 therapy[100]. Previous studies have reported that p53 mutations increase expression of PD-L1 *via*
562 decreased inhibition by miR-34[101] and that LUAD tumors with *KRAS-TP53* co-mutations have
563 higher expression of PD-L1 and improved response to anti-PD1 treatment compared to
564 *KRAS*^{mut}*TP53*^{WT} tumors[19,21,23,102]. The link between mutation status and immunotherapy response
565 has also been shown in the context of other mutations, where *KRAS*^{mut}*STK11*^{mut}*KEAP1*^{mut} LUAD
566 tumors demonstrate reduced response compared to *KRAS*^{mut}*STK11*^{WT}*KEAP1*^{WT} tumors[103]. In
567 agreement, our computational screen for immune checkpoint receptor-ligand pairs showed
568 enrichment in *PDCD1-CD274*, *CTLA4-CD86*, and *TIGIT-PVR* interactions in *TP53*^{mut} *vs*. *TP53*^{WT}
569 LUAD tumors; *CTLA4-CD86* and *TIGIT-PVR* interactions have not yet been studied in the context
570 of *TP53* mutations to the best of our knowledge. Increased expression of *TIGIT* in T cells and *PVR*
571 in malignant cell both contribute to the spatially-resolved enrichment of *TIGIT-PVR* co-
572 expression, which we observed in all 8 *TP53*^{mut} LUAD tumors in our cohort, as well as in bulk
573 RNA-seq data from TCGA and bulk proteomic data from CPTAC. Recent trials demonstrate
574 benefit to combining anti-PD1 ICI with chemotherapy as a neo-adjuvant in resectable LUAD[104].
575 Moreover, trials targeting TIGIT alone or in combination with PD1/PD-L1 in lung cancer have yet
576 to demonstrate substantial survival benefit, highlighting the importance of identifying subsets of

577 patient tumors that could benefit from therapeutic intervention (*e.g.* *TP53*^{mut}, PVR-expressing
578 tumors)[105]. Future studies will be needed to assess the predictive potential of *TP53* mutational
579 status for response to these and other therapies.
580
581 More generally, the methodology introduced in this study can be used to study the effect of *TP53*
582 or other mutations on the TME in lung cancer and other tumor types and to infer multicellular
583 pathways and communities to target for tumor suppressors in other contexts, where more effective
584 therapies are needed.
585
586

587 **METHODS**

588

589 **Human subjects**

590 Fresh solid primary tumor tissue was collected using an IRB-approved protocol (IRB protocol
591 98-063) at DFCI/BWH and MGH. Patients were all confirmed to be treatment-naïve and
592 consented to the study prior to collection. Immediately after resection, the tissue was reviewed
593 by the clinical pathology team and high-quality portions (determined based on tumor content,
594 necrosis, calcification, fat, and hemorrhage) were allocated for WES, scRNA-seq, spatial
595 transcriptomics, and mIF. The protocol was designed to reduce the time between surgical
596 resection, anatomic pathology review, placement in media, and processing at the Broad Institute.
597 Blood was drawn from the same patients and cryo-preserved at -80 °C for subsequent
598 processing. Patient demographic information, including age, sex, and smoking history, are
599 included in **Table S1**.

600

601 **Whole-exome sequencing sample processing**

602 DNA was extracted from fresh frozen lung cancer tissue embedded into OCT (TissueTek,
603 Sakura) and from PBMCs isolated from preserved patient-matched blood (AllPrep DNA/RNA
604 extraction kit, Qiagen). Library construction was performed as previously described[106]. In brief,
605 DNA input for shearing was diluted to a final concentration of 20–250 ng in 50 µl of solution.
606 Adapter ligation was performed using palindromic forked adapters (Integrated DNA
607 Technologies), containing unique dual-indexed molecular barcode sequences to improve
608 downstream pooling. End repair/A-tailing, adapter ligation, and library enrichment PCR were
609 carried out using a 96-reaction kit format (Kapa HyperPrep). During solid-phase reversible

610 immobilization (SPRI), a final elution volume of 30 µl was produced to maximize library
611 concentration, and vortexing was performed to maximize the effluent. Constructed libraries were
612 first pooled and hybridization and capture were performed (Illumina's Nextera Exome Kit) using
613 the recommended protocol from the manufacturer, using a skirted PCR plate to facilitate
614 automation (Agilent Bravo liquid handling system). Library pools then underwent qPCR
615 quantification and libraries were adjusted to a concentration of 2nM. DNA libraries were cluster-
616 amplified using exclusion amplification chemistry in patterned flow cells (Illumina) according to
617 the manufacturer's recommended protocol. Sequencing of flow cells was performed using
618 sequencing-by-synthesis chemistry and analyzed using RTA v.2.7.3 or later. Library pools were
619 sequenced on 76 cycle runs using two 8-cycle index reads across the appropriate number of lanes
620 to attain coverage for all libraries.
621
622 **Tissue processing, CD45 sorting, and scRNA-sequencing**
623 Tumor tissue resections were transferred in RPMI on ice from the operating room and washed in
624 cold PBS in the laboratory and transferred to 5ml Eppendorf tubes containing 3ml dissociation
625 mixture (NSCLC PDEC protocol[107]). Samples were minced in the Eppendorf tube using scissors
626 into pieces smaller than ~0.4 mm and incubated at 37 °C, rotating at ~14 rpm. for 10 min. Each
627 sample was pipetted 20 times with a 1ml pipette tip at room temperature and placed back into
628 incubation for 10 min. The sample was pipetted again 20 times using a 1 ml pipette tip,
629 transferred to a 1.7 ml Eppendorf tube and centrifuged at 300–580*g* for 4–7 min at 4 °C. The
630 pellet was resuspended in 200–500 µl of ACK (ammonium-chloride-potassium) RBC lysis buffer
631 (Thermo Fisher Scientific, A1049201) and incubated for 1 min on ice. Twice the volume of cold
632 PBS was added to stop the reaction, and cells were pelleted by a short centrifugation for 8 s at

633 4 °C, using the short spin setting and ramping up to 11,000*g*. If RBCs remained, the RBC lysis
634 step was performed up to two additional times. To assess cell count and viability, 5 µl of Trypan
635 blue (Thermo Fisher Scientific, cat. no. T10282) was mixed with 5 µl of the sample and loaded
636 onto an INCYTO C-Chip Disposable Hemocytometer, Neubauer Improved (VWR, cat. no.
637 82030-468). Depletion of $CD45^+$ cells in NSCLC samples was performed using CD45
638 MicroBeads (Miltenyi Biotec, cat. no. 130-045-801) according to the manufacturer's protocol.
639 The single-cell suspension was centrifuged at 500*g* for 4 min at 4 °C and the pellet was
640 resuspended in 80 µl of MACS buffer (PBS supplemented with 0.5% BSA, and 2 mM EDTA)
641 per 10^6 cells. MACS CD45 microbeads were added to the cell suspension (20 µl per 10 million
642 cells) and the cells were incubated on ice for 15 min. During incubation, a MS column was
643 attached to a MidiMACS separator and rinsed with 3 ml MACS buffer. After incubation, the
644 cells-bead conjugate was washed with 900 µl MACS buffer per 10 million cells. The cells were
645 then centrifuged at 500*g* for 4 min at 4 °C and the pellet was resuspended in 500 µl MACS
646 buffer. The cell suspension was transferred to the column and the effluent was collected
647 ($CD45^-$ fraction). The column was washed three times with 3 ml MACS buffer. The
648 $CD45^-$ fraction was centrifuged at 500*g* for 4 min at 4 °C. The pellet was resuspended in 50 µl of
649 0.4% BSA (Ambion, cat. no. AM2616) in PBS. Cells were re-counted and adjusted to a final
650 range of 200–2,000 cells per µl.
651
652 **NSCLC tissue dissociation protocol**
653 The digestion mix contained 2692 µl HBSS (Thermo Fisher Scientific, cat. no. 14170112),
654 187.5 µl of 20 mg ml^{-1} pronase (Sigma Aldrich, cat. no. 10165921001) diluted to a final
655 concentration of 1,250 µg ml^{-1}, 27.6 µl of 1 mg ml^{-1} elastase (Thermo Fisher Scientific, cat. no.

656 NC9301601) diluted to a final concentration of 9.2 μg ml^{-1}, 30 μl of 10 mg ml^{-1} DNase I (Sigma
657 Aldrich, cat. no. 11284932001) diluted to a final concentration of 100 μg ml^{-1}, 30 μl of
658 10 mg ml^{-1} Dispase (Sigma Aldrich, cat. no. 4942078001) diluted to a final concentration of
659 100 μg ml^{-1}, 30 μl of 150 mg ml^{-1} collagenase A (Sigma Aldrich, cat. no. 10103578001) diluted
660 to a final concentration of 1,500 μg ml^{-1} and 3 μl of 100 mg ml^{-1} collagenase IV (Thermo Fisher
661 Scientific, cat. no. NC9836075) diluted to a final concentration of 100 μg ml^{-1}.
662
663 **scRNA-seq library preparation and sequencing**
664 A total of 8,000 cells were loaded onto each channel of the Chromium Controller (10X
665 Genomics) to generate single-cell gel bead-in-emulsions (GEMs). Libraries were constructed
666 using the Chromium Single Cell 3′ Library & Gel Bead Kit v.2 or v.3 (PN-120237, 10X
667 Genomics): barcoded reverse transcription of RNA, cDNA amplification, fragmentation and
668 adapter and sample index attachment were all performed according to the manufacturer's
669 recommended protocol. Barcoded libraries from four 10x channels were pooled together and
670 sequenced on either one lane of an Illumina HiSeq X or one flow cell of a NextSeq, with paired
671 end reads accordingly: read 1, 26 nt; read 2, 55 nt; index 1, 8 nt; index 2, 0 nt.
672
673 **Tissue processing for Spatial Transcriptomics**
674 Fresh frozen lung cancer tissue samples were embedded into OCT (TissueTek, Sakura) and
675 shipped to the Broad Institute for cryosectioning and generation of hematoxylin and eosin-
676 stained (H&E) tissue sections. H&E-stained tissues were subject to pathology review to assess
677 tissue quality, structural preservation, cellular viability, tumor content, inflammation, fibrosis,
678 and necrosis. Samples were excluded based on small tissue size, low cellularity, or extensive

679 fibrosis. Regions of interest (ROIs) were then selected based on the criteria mentioned and
680 marked on H&E images for subsequent tissue sectioning and mounting on Visium slides. Visium
681 sectioning and processing were performed on the eight samples that passed QC. Tissues were
682 cryosectioned at 10 um thickness at -22°C and placed in the capture areas of Visium Tissue
683 Optimization slides (3000394, 10X Genomics) and Visium Spatial Gene Expression slides
684 (2000233, 10X Genomics). The tissue sections were adhered by warming the backside of the
685 slides and placed at -80°C for 1-3 days.
686
687 **Visium spatial gene expression library generation**
688 The tissue optimization sample slide and spatial gene expression slide were processed according
689 to the manufacturer's protocols
690 (CG000238_VisiumSpatialTissueOptimizationUserGuide_Rev_A.pdf and
691 CG000239_VisiumSpatialGeneExpression_UserGuide_Rev_C.pdf). In short, following tissue
692 methanol fixation and H&E staining, brightfield morphology images were obtained with a Zeiss
693 Axio microscope using the Metafer slide-scanning platform (Metasystems) at 10x resolution.
694 The images were joined together with the VSlide software (Metasystems) and exported as tiff
695 files. Optimal tissue permeabilization times were tested with eight different time points (0, 3, 6,
696 12, 18, 24, 30, 36 minutes) on the tissue optimization sample slide for one of the samples. 12
697 minutes of permeabilization was set as the optimal time point and used for the spatial gene
698 expression slide for all samples. The released RNAs from the tissue were reversely transcribed
699 into complimentary DNA by priming to the spatial barcoded primers on the glass in the presence
700 of template switching oligo. After second strand synthesis, a denaturation step released the
701 cDNAs, which were then amplified with 14-15 cycles of PCR amplification. Finally, indexed

702 sequencing-ready, spatial gene expression libraries were constructed according to the
703 manufacturer's protocol. The libraries were pooled together to generate >50,000 reads per spatial
704 spot and sequenced on Illumina NovaSeq sequencers with S1 and SP kits and the settings: read
705 1, 28 cycles; read 2, 90 cycles; index 1, 10 cycles; index 2, 10 cycles.
706

707 **Multiplex immunofluorescence**

708 Staining was performed on a BOND RX fully automated stainer (Leica Biosystems). FFPE tissue
709 sections of 5-µm thick were baked for 3 hours at 60°C before being loaded into the BOND RX.
710 Slides were deparaffinized (BOND DeWax Solution, Leica Biosystems, Cat. AR9590) and
711 rehydrated with multiple rounds of graded ethanol to deionized water. Antigen retrieval was
712 performed in BOND Epitope Retrieval Solution 1 (pH 6) or 2 (pH 9) (ER1, ER2, Leica
713 Biosystems, Cat. AR9961, AR9640) at 95°C. Deparaffinization, rehydration and antigen
714 retrieval are carried out by BOND RX for ~ 14 hours. Slides were then serially stained with
715 primary antibodies with an incubation time of 30 minutes per antibody. Anti-mouse and anti-
716 rabbit Opal Polymer Horseradish Peroxidase (Opal Polymer HRP Ms + Rb, Akoya Biosciences,
717 Cat. ARH1001EA) were then applied as a secondary label for 10 minutes. Signal for antibody
718 complexes was labeled and visualized by their corresponding Opal Fluorophore Reagents
719 (Akoya) by incubating the slides for 10 minutes. Opal Fluorophore 780 was paired with a TSA-
720 DIG amplification to ensure analyzable signal. Slides were incubated in Spectral DAPI solution
721 (Akoya) for 10 minutes, air dried and mounted with Prolong Diamond Anti-fade mounting
722 medium (Life Technologies, Cat. P36965) and stored in a light-proof box at 4℃ prior to
723 imaging.

724 Imaging was performed using the PhenoImager multispectral imaging platform (Akoya
725 Biosciences, Marlborough, MA). Each slide was scanned at 20x resolution as a whole-slide
726 image. These images were then accessed through Phenochart viewing software (Akoya
727 Biosciences) where 4-8 20x regions of interest (ROIs) were selected. After ROI selection was
728 approved by a pathologist (S.J.R.), the images were spectrally unmixed and analyzed using
729 Inform 2.6 (Akoya Biociences). Each analyzable ROI was segmented and quantified for
730 expression of each marker utilizing the Inform analysis tools.

731

732 **WES data processing**

733 The Picard pipeline (http://picard.sourceforge.net/) was used to align the tumor and normal
734 whole exome sequences to the hg19 reference human genome build and generate BAM files.
735 Somatic short variant discovery was performed using the Mutect 2 algorithm
736 (https://software.broadinstitute.org/gatk/documentation/tooldocs/current/org_broadinstitute_gatk
737 _tools_walkers_cancer_m2_MuTect2) from the GAT4K pipeline run on the Terra cloud-based
738 platform. Somatic mutations were annotated using Funcotator
739 (https://gatk.broadinstitute.org/hc/en-us/articles/360035889931-Funcotator-Information-and-
740 Tutorial). Copy number ratios for each exon was calculated by comparing mean exon coverage
741 with expected coverage based on a panel of normal samples and were then segmented for
742 downstream analysis. Maf and vcf files were generated as outputs, which were then subsequently
743 processed with maftools[108] and cnvkit[109] packages.

744

745 **Large-scale CNA inference and correlation**

746 The inferCNV package[110] (https://github.com/broadinstitute/inferCNV) was applied to infer
747 large-scale CNVs from scRNA-seq data. Normal lung epithelial cells from a separate dataset
748 were used as a reference. Raw UMI counts, a gene ordering file, and an annotation file
749 specifying normal and tumor cells and patient sample of origin were used as input. The function
750 infercnv::run() was performed with a cutoff of 0.1, cluster_by_groups = T, and HMM = T. CNV
751 values were predicted using a 6-state Hidden Markov Model (HMM) and were normalized to the
752 3rd quantile of all values per sample. Cells with > 0.7 normalized CNV values were assigned as
753 malignant cells. CNVs called from WES data were used as a positive control to assess the
754 accuracy of the inferCNV analysis using scRNA-seq data. Spearman correlation was run
755 between inferred CNVs from scRNA and from WES data across genes and plotted as a heatmap.
756

757 **scRNA-seq preprocessing**

758 CellRanger v3.1 was used to align 10X Chromium reads to the GRCh38 human genome
759 reference and generate barcode, gene, and count matrices. Preprocessed matrices were then
760 loaded into Seurat v4[111] in R v4.1.0 using the Read10X() function. Cells with less than 1000
761 unique molecular identifiers (UMIs), less than 400 detected genes (1000 genes for malignant
762 cells), or greater than 25% mitochondrial genes were excluded from further analyses.
763

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

765 The UMI counts matrices were batch-corrected by patient sample, using the Seurat v4
766 SCTransform() (using 2000 variable features and regressing out percentage of mitochondrial
767 genes), FindIntegrationAnchors(), and IntegrateData(), using the largest sample as the reference
768 dataset. PCA was then performed in the variable gene space and 15 principal components were

769 used for Louvain clustering and UMAP dimensionality reduction. Cell type markers for broad
770 cell classes were identified using the FindAllMarkers() function, using a maximum of 1,000 cells
771 per cluster. Cells were annotated based on expression of canonical markers that were established
772 by previous literature. Individual cell classes were then subset and subjected to another series of
773 integration, clustering, and annotation based on expression of subset-specific markers. Doublets
774 were removed based on expression of markers specific to more than one broad cell classes.
775

776 **scRNA-seq signature scoring**

777 To identify broad cell classes signatures, we used the FindAllMarkers() function in Seurat v4 and
778 generated gene sets using the most differentially expressed genes for each class, based on $\log_2$
779 fold change and Bonferroni-corrected *p*-values compared to the other clusters. We identified
780 malignant program signatures by performing anchor-based integration across patient samples
781 malignant cells followed by Louvain clustering at resolution 1.2 and using the FindAllMarkers()
782 function to derive the top 10 most differentially expressed genes from each cluster, based on $\log_2$
783 fold change and Bonferroni-corrected *p*-values. Signatures for normal lung epithelial cells were
784 derived from published healthy lung scRNA-seq data using the same method. The lists of broad
785 cell classes and cell subset marker genes, as well as final gene set signatures are provided in
786 **Table S2**. To define hallmark EMT and hypoxia scores, we used gene sets termed
787 "HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION" and
788 "HALLMARK_HYPOXIA" downloaded from MSigDB[112]. Healthy lung epithelial cell
789 signatures were derived using published scRNA-seq data from the human lung cell atlas[113], as
790 well as *a priori* knowledge: AT2 (*SFTPC*, *SFTPA2*, *SFTPA1*, *SFTPD*, *SFTPB*), AT1 (*AGER*,
791 *EMP2*, *CAV1*, *HOPX*), basal (*KRT5*, *KRT17*, *KRT15*, *S100A2*), neuroendocrine (*CALCA*,

792 *CHGA*, *GRP*, *SCG5*, *ASCL1*), goblet (*MUC5AC*, *MUC5B*, *S100P*), secretory (*BPIFB1*, *PIGR*,
793 *SCGB3A1*, *SCGB1A1*), ciliated (*TMEM190*, *CAPS*, *FOXJ1*, *TPPP3*, *FAM183A*), and ionocyte
794 (*ASCL3*, *ATP6V0B*, *FOXI1*, *CLCNKB*, *CFTR*). Signature scores were assigned per cell in
795 scRNA-seq data using the AddModuleScore() function in Seurat v4, which compares the average
796 expression of the genes in each gene set to the average expression of a control gene set of 100
797 randomly selected genes within the same expression bin.
798
799 **scRNA-seq differential expression analysis**
800 The data was first split into different cell subsets, and pseudobulk profiles were generated by
801 setting the identity of the objects to patient sample ID and running the AverageExpression()
802 function in Seurat v4. A Mann-Whitney Wilcoxon test was used to test for Bonferroni-corrected
803 significance in gene differential expression between groups. The average $\log_2$ fold change was
804 plotted using the ComplexHeatmap package in R, and asterisks were added for genes that were
805 significantly differentially expressed.
806
807 **Gene Set Enrichment Analysis**
808 Using the top 100 DEGs between conditions per cell subset or the top 100 specific markers per
809 subset, we performed gene-set enrichment analysis using the R package clusterProfiler[114] (v4.0),
810 focusing on GO terms, KEGG, Reactome, and Hallmark Cancer gene sets. Gene ratios and
811 adjusted *p*-values were visualized as dot plots, after applying BH correction and setting *p*-value
812 and *q*-value cutoffs of 0.05.
813
814 **Similarity to normal lung and K/KP mouse cells**

815 To identify transcriptomic similarities between cell subsets in NSCLC and the normal lung, the
816 Seurat v4 function TransferData() was used. The broad cell compartment object in the NSCLC
817 scRNA-seq dataset was assigned to the query parameter and the corresponding compartment in a
818 published normal lung scRNA-seq dataset[56] was assigned to the reference parameter, and the
819 remaining parameters were set to defaults. Each cell in the NSCLC object was assigned a
820 predicted identity based on the normal lung subset it was most similar to transcriptionally. A
821 confusion matrix was created, displaying the number of cells that overlapped between predicted
822 identities from the normal lung and annotations assigned *de novo*. A similar approach was used
823 to classify K/KP mice HPCS cells to their closest annotation in human tumor cells.
824
825 **Signaling Entropy calculation**
826 Entropy was inferred for each epithelial cell using the R package SCENT[64], using the default
827 protein-protein interaction (PPI) network ("net17Jan2016.m"). Raw UMI counts were used as
828 the scRNA-seq input. Entropy was calculated separately for malignant and non-malignant
829 epithelial cells. Quality control was performed using the scater package in R, and low-quality
830 cells based on library size, number of expressed genes, and mitochondrial gene proportion were
831 removed. Differential potency estimation was performed using the DoIntegPPI() function, which
832 generated entropy values per cell. The same method was used to calculate entropy values for
833 bulk RNA-seq data from TCGA and bulk proteomic data from CPTAC. Entropy values were
834 compared using the ggviolin() function using the ggpubr package in R.
835
836 **Cell-cell interaction analysis**

837 CellPhoneDB[115] (v2.1.7) was used to predict significant ligand-receptor interactions using
838 normalized raw counts and annotation of broad cell classes as input. The method was run with
839 default parameters separately for each sample ID to account for batch effects between samples.
840 Interactions with a *p*-value below 0.05 were classified as significant. The significant means
841 output was converted into a binary matrix based on presence or absence of a significant
842 interaction. Normalized enrichment difference (*ED*) of a specific interaction was calculated as
843 follows:

844 $$ED = n_A/N_A - n_B/N_B$$

845 where *nA* and *nB* represent the number of tumor samples with the interaction in groups A and B,
846 respectively, while *NA* and *NB* represent the total number of tumor samples in these groups.
847 Significance of enrichment for an interaction was calculated using a Fisher's exact test. Negative
848 $\log_{10}$ *p*-values and normalized enrichment differences for each ligand-receptor interaction and
849 cell pair combination were plotted using the ggplot2 package in R.
850

851 **scRNA-seq regulon analysis**

852 A scRNA-seq transcription factor (TF) gene regulatory network was constructed using
853 pySCENIC[73] (v0.11.2). Raw UMI counts were used as input and regulons were predicted using
854 gene module co-expression using the GRNBoost2 package. The number of regulons was pruned
855 using the feather ranking databases (hg19-500bp-upstream-7species.mc9nr, hg19-tss-centered-
856 10kb-10species.mc9nr, hg19-tss-centered-5kb-10species.mc9nr). Regulon activity, driving TFs,
857 and weights for individual target genes were predicted with the cisTarget() function in
858 pySCENIC. Regulon activity enrichment scores were predicted for each cellular subset using
859 AUCell and regulon specificity scores were used to identify regulatory networks specific to each

860 cell subset. Z-score-transformed regulon x subset matrices were visualized using the
861 ComplexHeatmap package in R.
862
863 **Additional scRNA-seq datasets**
864 Processed scRNA-seq data from an additional 5 studies (both published and unpublished) with
865 available *TP53* mutational status were assembled into a validation compendium, consisting of 15
866 *TP53*^{WT} and 7 *TP53*^{mut} tumor samples (ArrayExpress accessions E-MTAB-6149 and E-MTAB-
867 6653[26]; https://humantumoratlas.org/ (unpublished); GEO accession GSE123904[32]; BIG
868 accession HRA000154[55]; https://doi.org/10.24433/CO.0121060.v1[33]). *In-vivo* scRNA-seq data
869 for *KRAS*^{mut}*TP53*^{WT} and *KRAS*^{mut}*TP53*^{mut} mice (Smart-Seq2) was obtained from GEO accession
870 GSE152607[63]. *In-vitro* scRNA-seq data from lung cancer A549 cell lines with different induced
871 *TP53* variants was obtained from GEO accession GSE161824[61].
872
873 **Signature scoring of bulk expression profiles**
874 Bulk RNA-seq and somatic mutation data from TCGA LUAD and somatic mutations were
875 downloaded from cBioPortal[116] using the cgdsr package in R. Curated gene sets with highly
876 specific markers for individual cell types, subsets, or malignant programs were generated from
877 the scRNA-seq data as described above. The signature scores for these gene sets were defined as
878 the mean $\log_{10}$ expression of all the genes in a given gene set and were used to estimate the
879 relative frequency of cell types, cell subsets in bulk RNA-seq data from TCGA, or activity of
880 malignant cell programs. The same approach was used for calculating expression signatures in
881 bulk proteomic data from CPTAC[117], which was downloaded from the CPTAC Data
882 Portal https://cptac-data-portal.georgetown.edu/cptacPublic/. We confirm the reproducibility of

883 all our bulk deconvolution results using BayesPrism, a deconvolution method based on a
884 Bayesian framework that incorporates single-cell expression profiles[118].
885
886 **Survival analysis**
887 To predict the roles of individual genes and gene signatures on patient survival, TCGA LUAD
888 tumor samples were classified into high (top 25th percentile), medium (25th to 75th percentile),
889 and low (bottom 25th percentile); or high (top 50th percentile) and low (bottom 50th percentile)
890 expression groups. Kaplan-Meier (KM) curves were generated using the survival package in R,
891 and *p*-values were calculated using a log-rank test or bivariate Cox proportional-hazards
892 regression analysis to evaluate the significance of different model parameters (*e.g.*, expression
893 level of genes or malignant signatures, *TP53* mutation status) in relation to overall survival
894 (defined as time to death from any cause). The reported *p*-values from the Cox proportional-
895 hazards model reflect the significance of one specific covariate while adjusting for *TP53*
896 mutation status.
897
898 **Effects of co-mutations and *TP53* variant classes**
899 TCGA tumor samples were stratified into six mutually exclusive categories with sufficiently large
900 sample sizes for comparative analysis based on combinations of mutations in *KRAS*, *EGFR*, and
901 *TP53* (*KRAS*^{mut}, *EGFR*^{mut}, *TP53*^{mut}, *KRAS*^{mut}*TP53*^{mut}, *EGFR*^{mut}*TP53*^{mut}, and *KRAS*^{WT}*EGFR*^{WT}
902 *TP53*^{WT}). TCGA and CPTAC tumor samples were furthermore grouped into different mutually
903 exclusive missense variant classes based on either dominant-negative/loss of function (DNE,LOF)
904 status of mutational variants, or impact categories as described by prior studies[60,61]. Mann-

905 Whitney-Wilcoxon test *p*-values were computed for each group relative to the relevant wildtype
906 or control group.
907
908 **Multiple linear regression and ANOVA analysis of interaction effects**
909 Expression level of individual genes or gene signatures was predicted based on the mutation
910 status of four different genes (*EGFR*, *KRAS*, *STK11*, and *TP53*), and tumor mutational burden
911 (TMB) using the following model:
912 $$Y=\beta 0 + \beta 1X1 + \beta 2X2 + \beta 3X3 + \beta 4X4 + \beta 5X5 + \epsilon,$$
913 where Y represents the expression level, $\beta 0$ the y-intercept, $\beta 1$, $\beta 2$, $\beta 3$, $\beta 4$, and $\beta 5$ the regression
914 coefficients associated with each variable, $X1$, $X2$, $X3$, and $X4$ the mutation status of different
915 genes (*EGFR*, *KRAS*, *STK11*, and *TP53*), $X5$ the TMB, and ϵ the error term. For each mutation
916 variable Xi, a *p*-value was computed to test the null hypothesis that the regression coefficient βi
917 is equal to zero (indicating no effect). A significant *p*-value (≤ 0.05) suggests that the mutation
918 status of the respective gene is a significant predictor of Y (*i.e.,* there exists a significant
919 relationship between the mutation status and the expression level). We further conducted the
920 same analysis using a model containing an expanded set of covariates (mutation status of *EGFR*,
921 *KRAS-G12C*, *KRAS-G12D*, *KRAS-G12V*, *STK11*, *KEAP1*, *RBM10*, *PTPRD*, and *TP53* as well as
922 TMB). An analysis of variance (ANOVA) was conducted on the output of the regression
923 model's output, evaluating the interactions between *KRAS* and *TP53*, and between *EGFR* and
924 *TP53*. A significant interaction implies that the combined effect of two individual interactions is
925 not merely the sum of their individual effects. The output from the ANOVA provides *p*-values
926 for each interaction, where a significant *p*-value suggests that the effect of one mutation is altered
927 by the presence or absence of the other mutation (*i.e.,* not additive).

928

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

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

940

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

943 Tangram version 0.4.0, scanpy 1.8.1, and anndata 0.7.6 was used to map scRNA-seq data to
944 spatial locations on 10X Visium spatial transcriptomics samples as previously described[41]. UMI
945 counts for scRNA-seq and spatial transcriptomics data were converted into anndata objects and
946 pre-processed using standard scanpy functions[119]. The top 100 genes that best characterized each
947 broad cell classes were selected by sc.tl.rank_genes_groups(). This subset of genes was used
948 afterwards as training genes for Tangram alignment. To more accurately map single-cells onto
949 spatial transcriptomics spots, an estimate of cell density per spot was attained through watershed-
950 based nuclei segmentation of the hematoxylin and eosin-stained serial tissue section. Images

951 were first imported by squidpy 1.1.0[120], and smoothened using squidpy.im.process with a sigma
952 = 4, then segmented by squidpy.im.segment with method = "watershed", using parameters
953 channel = 0 to select the red color channel and thres = 120. A threshold of 120 was selected to
954 best separate nuclei from background, and sq.im.calculate_features was used to count the
955 number of segmented nuclei per spot after segmentation. Finally, mapping was performed by
956 tg.map_cells_to_space using mode = "constrained" and density_prior = # of nuclei per spot /
957 total # of nuclei to constrain alignment based on segmented nuclei density, returning
958 probabilities of spatial location on a per-cell basis, which were employed in later analysis. To
959 project individual gene expression onto spatial transcriptomics spots, tg.project_genes was used.
960 Cell type and subset proportion were predicted per spot by summing up cell x spot probabilities
961 and normalizing the total probability scores per spot to 1. Spatial colocalization was inferred
962 using Pearson correlation between predicted cell type and subset compositions, and additional
963 continuous attributes from the metadata across spots.
964
965 **NMF analysis for spatial transcriptomics data**
966 NMF analysis was adapted from cell2location[121] using cell subset proportions per spatial
967 transcriptomics spot as inputs, enabling analysis of spatially colocalized cellular programs. NMF
968 was trained five times for a range of k = 3 to 30 factors, and k = 14 was selected based on
969 stability of training across the five restarts and elucidating discrete spatially-delimited
970 compartments across spatial transcriptomics samples. Using scanpy, NMF weights were plotted
971 onto spatial transcriptomics data to investigate spatial-dependent patterning of NMF factors.
972 NMF weights per subset were visualized using the heatmap function in cell2location.
973

974 **Multiplex immunofluorescence analysis**

975 InForm-analyzed ROIs were processed using the Pythologist software[122]. Cell colocalization
976 involved three spatial categories: 1) direct contact, 2) adjacency within 6 nearest neighbors, and
977 3) proximity within 50um of PD-L1^{+}cytokeratin^{+} cells. Proportion of PD1^{+}CD8^{+} cells out of all
978 PD-L1^{+}cytokeratin^{+} neighboring cells were computed across these spatially defined categories
979 when at least 20 neighboring cells were present.

980

981

982 **REFERENCES**

983

984 1. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization
985 Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances
986 Since the 2004 Classification. J Thorac Oncol 2015;10(9):1243-1260. DOI:
987 10.1097/JTO.0000000000000630.
988 2. Cancer Genome Atlas Research N. Comprehensive genomic characterization of
989 squamous cell lung cancers. Nature 2012;489(7417):519-25. DOI: 10.1038/nature11404.
990 3. Cancer Genome Atlas Research N. Comprehensive molecular profiling of lung
991 adenocarcinoma. Nature 2014;511(7511):543-50. DOI: 10.1038/nature13385.
992 4. Hanna NH, Robinson AG, Temin S, et al. Therapy for Stage IV Non-Small-Cell Lung
993 Cancer With Driver Alterations: ASCO and OH (CCO) Joint Guideline Update. J Clin
994 Oncol 2021;39(9):1040-1091. DOI: 10.1200/JCO.20.03570.
995 5. Howlader N, Forjaz G, Mooradian MJ, et al. The Effect of Advances in Lung-Cancer
996 Treatment on Population Mortality. N Engl J Med 2020;383(7):640-649. DOI:
997 10.1056/NEJMoa1916623.
998 6. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell
999 lung cancer. Nature 2018;553(7689):446-454. DOI: 10.1038/nature25183.
1000 7. National Cancer Institute. Non-Small Cell Lung Cancer Treatment.
1001 (https://www.cancer.gov/types/lung/patient/non-small-cell-lung-treatment-pdq).
1002 8. Campbell JD, Alexandrov A, Kim J, et al. Distinct patterns of somatic genome alterations
1003 in lung adenocarcinomas and squamous cell carcinomas. Nat Genet 2016;48(6):607-16.
1004 DOI: 10.1038/ng.3564.

1005 9. Halvorsen AR, Silwal-Pandit L, Meza-Zepeda LA, et al. TP53 Mutation Spectrum in
1006 Smokers and Never Smoking Lung Cancer Patients. Front Genet 2016;7:85. DOI:
1007 10.3389/fgene.2016.00085.
1008 10. Kaubryte J, Lai AG. Pan-cancer prognostic genetic mutations and clinicopathological
1009 factors associated with survival outcomes: a systematic review. NPJ Precis Oncol
1010 2022;6(1):27. DOI: 10.1038/s41698-022-00269-5.
1011 11. Al Bakir M, Huebner A, Martinez-Ruiz C, et al. The evolution of non-small cell lung
1012 cancer metastases in TRACERx. Nature 2023;616(7957):534-542. DOI: 10.1038/s41586-
1013 023-05729-x.
1014 12. Van Egeren D, Kohli K, Warner JL, et al. Genomic analysis of early-stage lung cancer
1015 reveals a role for TP53 mutations in distant metastasis. Sci Rep 2022;12(1):19055. DOI:
1016 10.1038/s41598-022-21448-1.
1017 13. Duffy MJ, Synnott NC, O'Grady S, Crown J. Targeting p53 for the treatment of cancer.
1018 Semin Cancer Biol 2020. DOI: 10.1016/j.semcancer.2020.07.005.
1019 14. Bykov VJN, Eriksson SE, Bianchi J, Wiman KG. Targeting mutant p53 for efficient
1020 cancer therapy. Nat Rev Cancer 2018;18(2):89-102. DOI: 10.1038/nrc.2017.109.
1021 15. Zhu G, Pan C, Bei JX, et al. Mutant p53 in Cancer Progression and Targeted Therapies.
1022 Front Oncol 2020;10:595187. DOI: 10.3389/fonc.2020.595187.
1023 16. Uehara I, Tanaka N. Role of p53 in the Regulation of the Inflammatory Tumor
1024 Microenvironment and Tumor Suppression. Cancers 2018;10(7) (In English). DOI:
1025 ARTN 219
1026 10.3390/cancers10070219.

1027 17. Shi D, Jiang P. A Different Facet of p53 Function: Regulation of Immunity and
1028 Inflammation During Tumor Development. Front Cell Dev Biol 2021;9:762651. DOI:
1029 10.3389/fcell.2021.762651.
1030 18. Hodis E, Torlai Triglia E, Kwon JYH, et al. Stepwise-edited, human melanoma models
1031 reveal mutations' effect on tumor and microenvironment. Science
1032 2022;376(6592):eabi8175. DOI: 10.1126/science.abi8175.
1033 19. Skoulidis F, Goldberg ME, Greenawalt DM, et al. STK11/LKB1 Mutations and PD-1
1034 Inhibitor Resistance in KRAS-Mutant Lung Adenocarcinoma. Cancer Discov
1035 2018;8(7):822-835. DOI: 10.1158/2159-8290.CD-18-0099.
1036 20. Biton J, Mansuet-Lupo A, Pecuchet N, et al. TP53, STK11, and EGFR Mutations Predict
1037 Tumor Immune Profile and the Response to Anti-PD-1 in Lung Adenocarcinoma. Clin
1038 Cancer Res 2018;24(22):5710-5723. DOI: 10.1158/1078-0432.CCR-18-0163.
1039 21. Dong ZY, Zhong WZ, Zhang XC, et al. Potential Predictive Value of TP53 and KRAS
1040 Mutation Status for Response to PD-1 Blockade Immunotherapy in Lung
1041 Adenocarcinoma. Clin Cancer Res 2017;23(12):3012-3024. DOI: 10.1158/1078-
1042 0432.CCR-16-2554.
1043 22. Leader AM, Grout JA, Maier BB, et al. Single-cell analysis of human non-small cell lung
1044 cancer lesions refines tumor classification and patient stratification. Cancer Cell
1045 2021;39(12):1594-1609 e12. DOI: 10.1016/j.ccell.2021.10.009.
1046 23. Davis AP, Cooper WA, Boyer M, Lee JH, Pavlakis N, Kao SC. Efficacy of
1047 immunotherapy in KRAS-mutant non-small-cell lung cancer with comutations.
1048 Immunotherapy-Uk 2021;13(11):941-952. DOI: 10.2217/imt-2021-0090.

1049 24. Scalera S, Mazzotta M, Corleone G, et al. KEAP1 and TP53 Frame Genomic,
1050 Evolutionary, and Immunologic Subtypes of Lung Adenocarcinoma With Different
1051 Sensitivity to Immunotherapy. J Thorac Oncol 2021;16(12):2065-2077. DOI:
1052 10.1016/j.jtho.2021.08.010.
1053 25. Lavin Y, Kobayashi S, Leader A, et al. Innate Immune Landscape in Early Lung
1054 Adenocarcinoma by Paired Single-Cell Analyses. Cell 2017;169(4):750-765 e17. DOI:
1055 10.1016/j.cell.2017.04.014.
1056 26. Lambrechts D, Wauters E, Boeckx B, et al. Phenotype molding of stromal cells in the
1057 lung tumor microenvironment. Nat Med 2018;24(8):1277-1289. DOI: 10.1038/s41591-
1058 018-0096-5.
1059 27. Zilionis R, Engblom C, Pfirschke C, et al. Single-Cell Transcriptomics of Human and
1060 Mouse Lung Cancers Reveals Conserved Myeloid Populations across Individuals and
1061 Species. Immunity 2019;50(5):1317-1334 e10. DOI: 10.1016/j.immuni.2019.03.009.
1062 28. Zhang L, Zhang Y, Wang C, et al. Integrated single-cell RNA sequencing analysis
1063 reveals distinct cellular and transcriptional modules associated with survival in lung
1064 cancer. Signal Transduct Target Ther 2022;7(1):9. DOI: 10.1038/s41392-021-00824-9.
1065 29. Wang Z, Li Z, Zhou K, et al. Deciphering cell lineage specification of human lung
1066 adenocarcinoma with single-cell RNA sequencing. Nat Commun 2021;12(1):6500. DOI:
1067 10.1038/s41467-021-26770-2.
1068 30. Wang Y, Liu B, Min Q, et al. Spatial transcriptomics delineates molecular features and
1069 cellular plasticity in lung adenocarcinoma progression. Cell Discov 2023;9(1):96. DOI:
1070 10.1038/s41421-023-00591-7.

1071 31. Kim N, Kim HK, Lee K, et al. Single-cell RNA sequencing demonstrates the molecular
1072 and cellular reprogramming of metastatic lung adenocarcinoma. Nat Commun
1073 2020;11(1):2285. DOI: 10.1038/s41467-020-16164-1.
1074 32. Laughney AM, Hu J, Campbell NR, et al. Regenerative lineages and immune-mediated
1075 pruning in lung cancer metastasis. Nat Med 2020;26(2):259-269. DOI: 10.1038/s41591-
1076 019-0750-6.
1077 33. Bischoff P, Trinks A, Obermayer B, et al. Single-cell RNA sequencing reveals distinct
1078 tumor microenvironmental patterns in lung adenocarcinoma. Oncogene
1079 2021;40(50):6748-6758. DOI: 10.1038/s41388-021-02054-3.
1080 34. He D, Wang D, Lu P, et al. Single-cell RNA sequencing reveals heterogeneous tumor and
1081 immune cell populations in early-stage lung adenocarcinomas harboring EGFR
1082 mutations. Oncogene 2021;40(2):355-368. DOI: 10.1038/s41388-020-01528-0.
1083 35. Alessi JV, Elkrief A, Ricciuti B, et al. Clinicopathologic and Genomic Factors Impacting
1084 Efficacy of First-Line Chemoimmunotherapy in Advanced NSCLC. J Thorac Oncol
1085 2023;18(6):731-743. DOI: 10.1016/j.jtho.2023.01.091.
1086 36. Camidge DR, Kim HR, Ahn MJ, et al. Brigatinib Versus Crizotinib in ALK Inhibitor-
1087 Naive Advanced ALK-Positive NSCLC: Final Results of Phase 3 ALTA-1L Trial. J
1088 Thorac Oncol 2021;16(12):2091-2108. DOI: 10.1016/j.jtho.2021.07.035.
1089 37. Stuart T, Butler A, Hoffman P, et al. Comprehensive Integration of Single-Cell Data. Cell
1090 2019;177(7):1888-1902 e21. DOI: 10.1016/j.cell.2019.05.031.
1091 38. Wang P, Wang F, He H, et al. TP53 and CDKN2A mutations in patients with early-stage
1092 lung squamous cell carcinoma: an analysis of the correlations and prognostic outcomes.
1093 Ann Transl Med 2021;9(16):1330. DOI: 10.21037/atm-21-3709.

1094 39. Dudley AC. Tumor endothelial cells. Cold Spring Harb Perspect Med
1095 2012;2(3):a006536. DOI: 10.1101/cshperspect.a006536.
1096 40. Cooke VG, LeBleu VS, Keskin D, et al. Pericyte depletion results in hypoxia-associated
1097 epithelial-to-mesenchymal transition and metastasis mediated by met signaling pathway.
1098 Cancer Cell 2012;21(1):66-81. DOI: 10.1016/j.ccr.2011.11.024.
1099 41. Biancalani T, Scalia G, Buffoni L, et al. Deep learning and alignment of spatially
1100 resolved single-cell transcriptomes with Tangram. Nat Methods 2021;18(11):1352-1362.
1101 DOI: 10.1038/s41592-021-01264-7.
1102 42. Tirosh I, Izar B, Prakadan SM, et al. Dissecting the multicellular ecosystem of metastatic
1103 melanoma by single-cell RNA-seq. Science 2016;352(6282):189-96. DOI:
1104 10.1126/science.aad0501.
1105 43. Patel AP, Tirosh I, Trombetta JJ, et al. Single-cell RNA-seq highlights intratumoral
1106 heterogeneity in primary glioblastoma. Science 2014;344(6190):1396-401. DOI:
1107 10.1126/science.1254257.
1108 44. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell
1109 transcriptomic data across different conditions, technologies, and species. Nature
1110 biotechnology 2018;36(5):411.
1111 45. Cangul H. Hypoxia upregulates the expression of the NDRG1 gene leading to its
1112 overexpression in various human cancers. BMC Genet 2004;5:27. DOI: 10.1186/1471-
1113 2156-5-27.
1114 46. Zhang JZ, Behrooz A, Ismail-Beigi F. Regulation of glucose transport by hypoxia. Am J
1115 Kidney Dis 1999;34(1):189-202. DOI: 10.1016/s0272-6386(99)70131-9.

1116 47. Cui XG, Han ZT, He SH, et al. HIF1/2alpha mediates hypoxia-induced LDHA
1117 expression in human pancreatic cancer cells. Oncotarget 2017;8(15):24840-24852. DOI:
1118 10.18632/oncotarget.15266.
1119 48. Zhang Z, Deng X, Liu Y, Liu Y, Sun L, Chen F. PKM2, function and expression and
1120 regulation. Cell Biosci 2019;9:52. DOI: 10.1186/s13578-019-0317-8.
1121 49. Puram SV, Tirosh I, Parikh AS, et al. Single-Cell Transcriptomic Analysis of Primary
1122 and Metastatic Tumor Ecosystems in Head and Neck Cancer. Cell 2017;171(7):1611-+.
1123 (In English). DOI: 10.1016/j.cell.2017.10.044.
1124 50. Baron M, Tagore M, Hunter MV, et al. The Stress-Like Cancer Cell State Is a Consistent
1125 Component of Tumorigenesis. Cell Syst 2020;11(5):536-546 e7. DOI:
1126 10.1016/j.cels.2020.08.018.
1127 51. Kinker GS, Greenwald AC, Tal R, et al. Pan-cancer single-cell RNA-seq identifies
1128 recurring programs of cellular heterogeneity. Nat Genet 2020;52(11):1208-1218. DOI:
1129 10.1038/s41588-020-00726-6.
1130 52. Min JW, Kim WJ, Han JA, et al. Identification of Distinct Tumor Subpopulations in
1131 Lung Adenocarcinoma via Single-Cell RNA-seq. PLoS One 2015;10(8):e0135817. DOI:
1132 10.1371/journal.pone.0135817.
1133 53. Tang H, Wang S, Xiao G, et al. Comprehensive evaluation of published gene expression
1134 prognostic signatures for biomarker-based lung cancer clinical studies. Ann Oncol
1135 2017;28(4):733-740. DOI: 10.1093/annonc/mdw683.
1136 54. Nagy A, Munkacsy G, Gyorffy B. Pancancer survival analysis of cancer hallmark genes.
1137 Sci Rep 2021;11(1):6047. DOI: 10.1038/s41598-021-84787-5.

1138 55. Xing X, Yang F, Huang Q, et al. Decoding the multicellular ecosystem of lung
1139 adenocarcinoma manifested as pulmonary subsolid nodules by single-cell RNA
1140 sequencing. Sci Adv 2021;7(5). DOI: 10.1126/sciadv.abd9738.
1141 56. Travaglini KJ, Nabhan AN, Penland L, et al. A molecular cell atlas of the human lung
1142 from single-cell RNA sequencing. Nature 2020;587(7835):619-625. DOI:
1143 10.1038/s41586-020-2922-4.
1144 57. Ferone G, Lee MC, Sage J, Berns A. Cells of origin of lung cancers: lessons from mouse
1145 studies. Genes Dev 2020;34(15-16):1017-1032. DOI: 10.1101/gad.338228.120.
1146 58. LaFave LM, Kartha VK, Ma S, et al. Epigenomic State Transitions Characterize Tumor
1147 Progression in Mouse Lung Adenocarcinoma. Cancer Cell 2020;38(2):212-228 e13. DOI:
1148 10.1016/j.ccell.2020.06.006.
1149 59. Yang L, He YT, Dong S, et al. Single-cell transcriptome analysis revealed a suppressive
1150 tumor immune microenvironment in EGFR mutant lung adenocarcinoma. J Immunother
1151 Cancer 2022;10(2). DOI: 10.1136/jitc-2021-003534.
1152 60. Giacomelli AO, Yang X, Lintner RE, et al. Mutational processes shape the landscape of
1153 TP53 mutations in human cancer. Nat Genet 2018;50(10):1381-1387. DOI:
1154 10.1038/s41588-018-0204-y.
1155 61. Ursu O, Neal JT, Shea E, et al. Massively parallel phenotyping of coding variants in
1156 cancer with Perturb-seq. Nat Biotechnol 2022;40(6):896-905. DOI: 10.1038/s41587-021-
1157 01160-7.
1158 62. Baslan T, Morris JPt, Zhao Z, et al. Ordered and deterministic cancer genome evolution
1159 after p53 loss. Nature 2022;608(7924):795-802. DOI: 10.1038/s41586-022-05082-5.

1160 63. Marjanovic ND, Hofree M, Chan JE, et al. Emergence of a High-Plasticity Cell State
1161 during Lung Cancer Evolution. Cancer Cell 2020;38(2):229-+. (In English). DOI:
1162 10.1016/j.ccell.2020.06.012.
1163 64. Teschendorff AE, Enver T. Single-cell entropy for accurate estimation of differentiation
1164 potency from a cell's transcriptome. Nat Commun 2017;8:15599. DOI:
1165 10.1038/ncomms15599.
1166 65. Schupp JC, Adams TS, Cosme C, Jr., et al. Integrated Single-Cell Atlas of Endothelial
1167 Cells of the Human Lung. Circulation 2021;144(4):286-302. DOI:
1168 10.1161/CIRCULATIONAHA.120.052318.
1169 66. Gillich A, Zhang F, Farmer CG, et al. Capillary cell-type specialization in the alveolus.
1170 Nature 2020;586(7831):785-789. DOI: 10.1038/s41586-020-2822-7.
1171 67. Muz B, de la Puente P, Azab F, Azab AK. The role of hypoxia in cancer progression,
1172 angiogenesis, metastasis, and resistance to therapy. Hypoxia (Auckl) 2015;3:83-92. DOI:
1173 10.2147/HP.S93413.
1174 68. Acevedo LM, Barillas S, Weis SM, Gothert JR, Cheresh DA. Semaphorin 3A suppresses
1175 VEGF-mediated angiogenesis yet acts as a vascular permeability factor. Blood
1176 2008;111(5):2674-80. DOI: 10.1182/blood-2007-08-110205.
1177 69. Guttmann-Raviv N, Shraga-Heled N, Varshavsky A, Guimaraes-Sternberg C, Kessler O,
1178 Neufeld G. Semaphorin-3A and semaphorin-3F work together to repel endothelial cells
1179 and to inhibit their survival by induction of apoptosis. J Biol Chem 2007;282(36):26294-
1180 305. DOI: 10.1074/jbc.M609711200.
1181 70. Pasquale EB. Eph receptors and ephrins in cancer: bidirectional signalling and beyond.
1182 Nat Rev Cancer 2010;10(3):165-80. DOI: 10.1038/nrc2806.

1183 71. Irvine AF, Waise S, Green EW, Stuart B, Thomas GJ. Characterising cancer-associated
1184 fibroblast heterogeneity in non-small cell lung cancer: a systematic review and meta-
1185 analysis. Sci Rep 2021;11(1):3727. DOI: 10.1038/s41598-021-81796-2.
1186 72. Buechler MB, Pradhan RN, Krishnamurty AT, et al. Cross-tissue organization of the
1187 fibroblast lineage. Nature 2021;593(7860):575-579. DOI: 10.1038/s41586-021-03549-5.
1188 73. Van de Sande B, Flerin C, Davie K, et al. A scalable SCENIC workflow for single-cell
1189 gene regulatory network analysis. Nat Protoc 2020;15(7):2247-2276. DOI:
1190 10.1038/s41596-020-0336-2.
1191 74. Lee KW, Yeo SY, Sung CO, Kim SH. Twist1 is a key regulator of cancer-associated
1192 fibroblasts. Cancer Res 2015;75(1):73-85. DOI: 10.1158/0008-5472.CAN-14-0350.
1193 75. De Wever O, Demetter P, Mareel M, Bracke M. Stromal myofibroblasts are drivers of
1194 invasive cancer growth. Int J Cancer 2008;123(10):2229-38. DOI: 10.1002/ijc.23925.
1195 76. Engblom C, Pfirschke C, Pittet MJ. The role of myeloid cells in cancer therapies. Nat
1196 Rev Cancer 2016;16(7):447-62. DOI: 10.1038/nrc.2016.54.
1197 77. Dong B, Wu C, Huang L, Qi Y. Macrophage-Related SPP1 as a Potential Biomarker for
1198 Early Lymph Node Metastasis in Lung Adenocarcinoma. Front Cell Dev Biol
1199 2021;9:739358. DOI: 10.3389/fcell.2021.739358.
1200 78. Heimberg G, Kuo T, DePianto D, et al. Scalable querying of human cell atlases via a
1201 foundational model reveals commonalities across fibrosis-associated macrophages.
1202 bioRxiv 2023:2023.07.18.549537. DOI: 10.1101/2023.07.18.549537.
1203 79. Casazza A, Laoui D, Wenes M, et al. Impeding macrophage entry into hypoxic tumor
1204 areas by Sema3A/Nrp1 signaling blockade inhibits angiogenesis and restores antitumor
1205 immunity. Cancer Cell 2013;24(6):695-709. DOI: 10.1016/j.ccr.2013.11.007.

1206 80. Zhang F, Wang H, Wang X, et al. TGF-beta induces M2-like macrophage polarization
1207 via SNAIL-mediated suppression of a pro-inflammatory phenotype. Oncotarget
1208 2016;7(32):52294-52306. DOI: 10.18632/oncotarget.10561.
1209 81. van der Leun AM, Thommen DS, Schumacher TN. CD8(+) T cell states in human
1210 cancer: insights from single-cell analysis. Nat Rev Cancer 2020;20(4):218-232. DOI:
1211 10.1038/s41568-019-0235-4.
1212 82. Litchfield K, Reading JL, Puttick C, et al. Meta-analysis of tumor- and T cell-intrinsic
1213 mechanisms of sensitization to checkpoint inhibition. Cell 2021;184(3):596-614 e14.
1214 DOI: 10.1016/j.cell.2021.01.002.
1215 83. Zhou X, Du J, Liu C, et al. A Pan-Cancer Analysis of CD161, a Potential New Immune
1216 Checkpoint. Front Immunol 2021;12:688215. DOI: 10.3389/fimmu.2021.688215.
1217 84. Hollern DP, Xu N, Thennavan A, et al. B Cells and T Follicular Helper Cells Mediate
1218 Response to Checkpoint Inhibitors in High Mutation Burden Mouse Models of Breast
1219 Cancer. Cell 2019;179(5):1191-1206 e21. DOI: 10.1016/j.cell.2019.10.028.
1220 85. Gide TN, Quek C, Menzies AM, et al. Distinct Immune Cell Populations Define
1221 Response to Anti-PD-1 Monotherapy and Anti-PD-1/Anti-CTLA-4 Combined Therapy.
1222 Cancer Cell 2019;35(2):238-255 e6. DOI: 10.1016/j.ccell.2019.01.003.
1223 86. Thommen DS, Schumacher TN. T Cell Dysfunction in Cancer. Cancer Cell
1224 2018;33(4):547-562. DOI: 10.1016/j.ccell.2018.03.012.
1225 87. Upadhaya S, Hubbard-Lucey VM, Yu JX. Immuno-oncology drug development forges
1226 on despite COVID-19. Nature reviews Drug Discovery 2020.

1227 88. Chen J, Tan Y, Sun F, et al. Single-cell transcriptome and antigen-immunoglobin analysis
1228 reveals the diversity of B cells in non-small cell lung cancer. Genome Biol
1229 2020;21(1):152. DOI: 10.1186/s13059-020-02064-6.
1230 89. Yu Y, Xiao CH, Tan LD, Wang QS, Li XQ, Feng YM. Cancer-associated fibroblasts
1231 induce epithelial-mesenchymal transition of breast cancer cells through paracrine TGF-
1232 beta signalling. Br J Cancer 2014;110(3):724-32. DOI: 10.1038/bjc.2013.768.
1233 90. Bonde AK, Tischler V, Kumar S, Soltermann A, Schwendener RA. Intratumoral
1234 macrophages contribute to epithelial-mesenchymal transition in solid tumors. Bmc
1235 Cancer 2012;12:35. DOI: 10.1186/1471-2407-12-35.
1236 91. Ventura A, Kirsch DG, McLaughlin ME, et al. Restoration of p53 function leads to
1237 tumour regression in vivo. Nature 2007;445(7128):661-5. DOI: 10.1038/nature05541.
1238 92. Tam SY, Wu VWC, Law HKW. Hypoxia-Induced Epithelial-Mesenchymal Transition in
1239 Cancers: HIF-1alpha and Beyond. Front Oncol 2020;10:486. DOI:
1240 10.3389/fonc.2020.00486.
1241 93. Joseph JP, Harishankar MK, Pillai AA, Devi A. Hypoxia induced EMT: A review on the
1242 mechanism of tumor progression and metastasis in OSCC. Oral Oncol 2018;80:23-32.
1243 DOI: 10.1016/j.oraloncology.2018.03.004.
1244 94. Graeber TG, Osmanian C, Jacks T, et al. Hypoxia-mediated selection of cells with
1245 diminished apoptotic potential in solid tumours. Nature 1996;379(6560):88-91. DOI:
1246 10.1038/379088a0.
1247 95. Tang Q, Su Z, Gu W, Rustgi AK. Mutant p53 on the Path to Metastasis. Trends Cancer
1248 2020;6(1):62-73. DOI: 10.1016/j.trecan.2019.11.004.

1249 96. Capaci V, Mantovani F, Del Sal G. Amplifying Tumor-Stroma Communication: An
1250 Emerging Oncogenic Function of Mutant p53. Front Oncol 2020;10:614230. DOI:
1251 10.3389/fonc.2020.614230.
1252 97. Pelka K, Hofree M, Chen JH, et al. Spatially organized multicellular immune hubs in
1253 human colorectal cancer. Cell 2021;184(18):4734-4752 e20. DOI:
1254 10.1016/j.cell.2021.08.003.
1255 98. Chen JH, Nieman LT, Spurrell M, et al. Spatial analysis of human lung cancer reveals
1256 organized immune hubs enriched for stem-like CD8 T cells and associated with
1257 immunotherapy response. bioRxiv 2023. DOI: 10.1101/2023.04.04.535379.
1258 99. Li L, Li M, Wang, X. Cancer type-dependent correlations between TP53 mutations and
1259 antitumor immunity. DNA Repair 2020;88. DOI:
1260 https://doi.org/10.1016/j.dnarep.2020.102785.
1261 100. Assoun S, Theou-Anton N, Nguenang M, et al. Association of TP53 mutations with
1262 response and longer survival under immune checkpoint inhibitors in advanced non-small-
1263 cell lung cancer. Lung Cancer 2019;132:65-71. DOI: 10.1016/j.lungcan.2019.04.005.
1264 101. Cortez MA, Ivan C, Valdecanas D, et al. PDL1 Regulation by p53 via miR-34. J Natl
1265 Cancer Inst 2016;108(1). DOI: 10.1093/jnci/djv303.
1266 102. Skoulidis F, Byers LA, Diao L, et al. Co-occurring genomic alterations define major
1267 subsets of KRAS-mutant lung adenocarcinoma with distinct biology, immune profiles,
1268 and therapeutic vulnerabilities. Cancer Discov 2015;5(8):860-77. DOI: 10.1158/2159-
1269 8290.CD-14-1236.
1270 103. Ricciuti B, Arbour KC, Lin JJ, et al. Diminished Efficacy of Programmed Death-
1271 (Ligand)1 Inhibition in STK11- and KEAP1-Mutant Lung Adenocarcinoma Is Affected

1272 by KRAS Mutation Status. J Thorac Oncol 2022;17(3):399-410. DOI:
1273 10.1016/j.jtho.2021.10.013.
1274 104. Forde PM, Spicer J, Lu S, et al. Neoadjuvant Nivolumab plus Chemotherapy in
1275 Resectable Lung Cancer. N Engl J Med 2022. DOI: 10.1056/NEJMoa2202170.
1276 105. Rousseau A, Parisi C, Barlesi F. Anti-TIGIT therapies for solid tumors: a systematic
1277 review. ESMO Open 2023;8(2):101184. DOI: 10.1016/j.esmoop.2023.101184.
1278 106. Jerby-Arnon L, Neftel C, Shore ME, et al. Opposing immune and genetic mechanisms
1279 shape oncogenic programs in synovial sarcoma. Nat Med 2021;27(2):289-300. DOI:
1280 10.1038/s41591-020-01212-6.
1281 107. Slyper M, Porter CBM, Ashenberg O, et al. A single-cell and single-nucleus RNA-Seq
1282 toolbox for fresh and frozen human tumors. Nat Med 2020;26(5):792-802. DOI:
1283 10.1038/s41591-020-0844-1.
1284 108. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and
1285 comprehensive analysis of somatic variants in cancer. Genome Res 2018;28(11):1747-
1286 1756. DOI: 10.1101/gr.239244.118.
1287 109. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number
1288 Detection and Visualization from Targeted DNA Sequencing. PLoS Comput Biol
1289 2016;12(4):e1004873. DOI: 10.1371/journal.pcbi.1004873.
1290 110. inferCNV of the Trinity CTAT Project. (https://github.com/broadinstitute/inferCNV).
1291 111. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell
1292 data. Cell 2021;184(13):3573-3587 e29. DOI: 10.1016/j.cell.2021.04.048.

1293 112. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The
1294 Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst
1295 2015;1(6):417-425. DOI: 10.1016/j.cels.2015.12.004.
1296 113. Sikkema L, Ramirez-Suastegui C, Strobl DC, et al. An integrated cell atlas of the lung in
1297 health and disease. Nat Med 2023;29(6):1563-1577. DOI: 10.1038/s41591-023-02327-2.
1298 114. Wu T, Hu E, Xu S, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting
1299 omics data. Innovation (N Y) 2021;2(3):100141. DOI: 10.1016/j.xinn.2021.100141.
1300 115. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring
1301 cell-cell communication from combined expression of multi-subunit ligand-receptor
1302 complexes. Nat Protoc 2020;15(4):1484-1506. DOI: 10.1038/s41596-020-0292-x.
1303 116. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics
1304 and clinical profiles using the cBioPortal. Sci Signal 2013;6(269):pl1. DOI:
1305 10.1126/scisignal.2004088.
1306 117. Gillette MA, Satpathy S, Cao S, et al. Proteogenomic Characterization Reveals
1307 Therapeutic Vulnerabilities in Lung Adenocarcinoma. Cell 2020;182(1):200-225 e35.
1308 DOI: 10.1016/j.cell.2020.06.013.
1309 118. Chu T, Danko CG. Bayesian cell-type deconvolution and gene expression inference
1310 reveals tumor-microenvironment interactions. 2020. DOI: 10.1101/2020.01.07.897900.
1311 119. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data
1312 analysis. Genome Biol 2018;19(1):15. DOI: 10.1186/s13059-017-1382-0.
1313 120. Palla G, Spitzer H, Klein M, et al. Squidpy: a scalable framework for spatial omics
1314 analysis. Nat Methods 2022;19(2):171-178. DOI: 10.1038/s41592-021-01358-2.

1315 121. Kleshchevnikov V, Shmatko A, Dann E, et al. Cell2location maps fine-grained cell types
1316 in spatial transcriptomics. Nat Biotechnol 2022. DOI: 10.1038/s41587-021-01139-4.
1317 122. Griffin GK, Weirather JL, Roemer MGM, et al. Spatial signatures identify immune
1318 escape via PD-1 as a defining feature of T-cell/histiocyte-rich large B-cell lymphoma.
1319 Blood 2021;137(10):1353-1364. DOI: 10.1182/blood.2020006464.
1320
1321

1322 **ACKNOWLEDGMENTS**

1323 We thank Leslie Gaffney for extensive help in editing and graphical design of the figures in the
1324 manuscript, the Broad Genomics Platform, Broad Flow Cytometry Facility, Pathology and Surgery
1325 Departments at MGH and BWH; members of the Bueno, Rodig, Hata, Regev, Johnson, and
1326 Tsankov labs. The project is part of the Human Tumor Atlas Pilot Project (HTAPP) consortium
1327 and National Institutes of Health (NIH) HTAN (Human Tumor Atlas Network) consortium paper
1328 package. Work was supported by the Klarman Cell Observatory (to O.R.-R. and A. Regev) and
1329 also funded in part by federal funds from the National Cancer Institute (NCI), NIH Task Order
1330 HHSN261100039 under Contract HHSN261201500003, and U2CCA233195. A.M.T. was
1331 supported by Department of the Army Lung Cancer Research Program Career Development
1332 Award W81XWH2210079, American Cancer Society Research Scholar Grant RSG-23-1039063-
1333 01-MM, and ISMMS seed funding. W.Z. and B.Y.S. were supported by NIH grant
1334 2T32GM007280-41 as part of the ISMMS Medical Scientist Training Program. The graphical
1335 summary was created with Biorender.com.

Figure 1

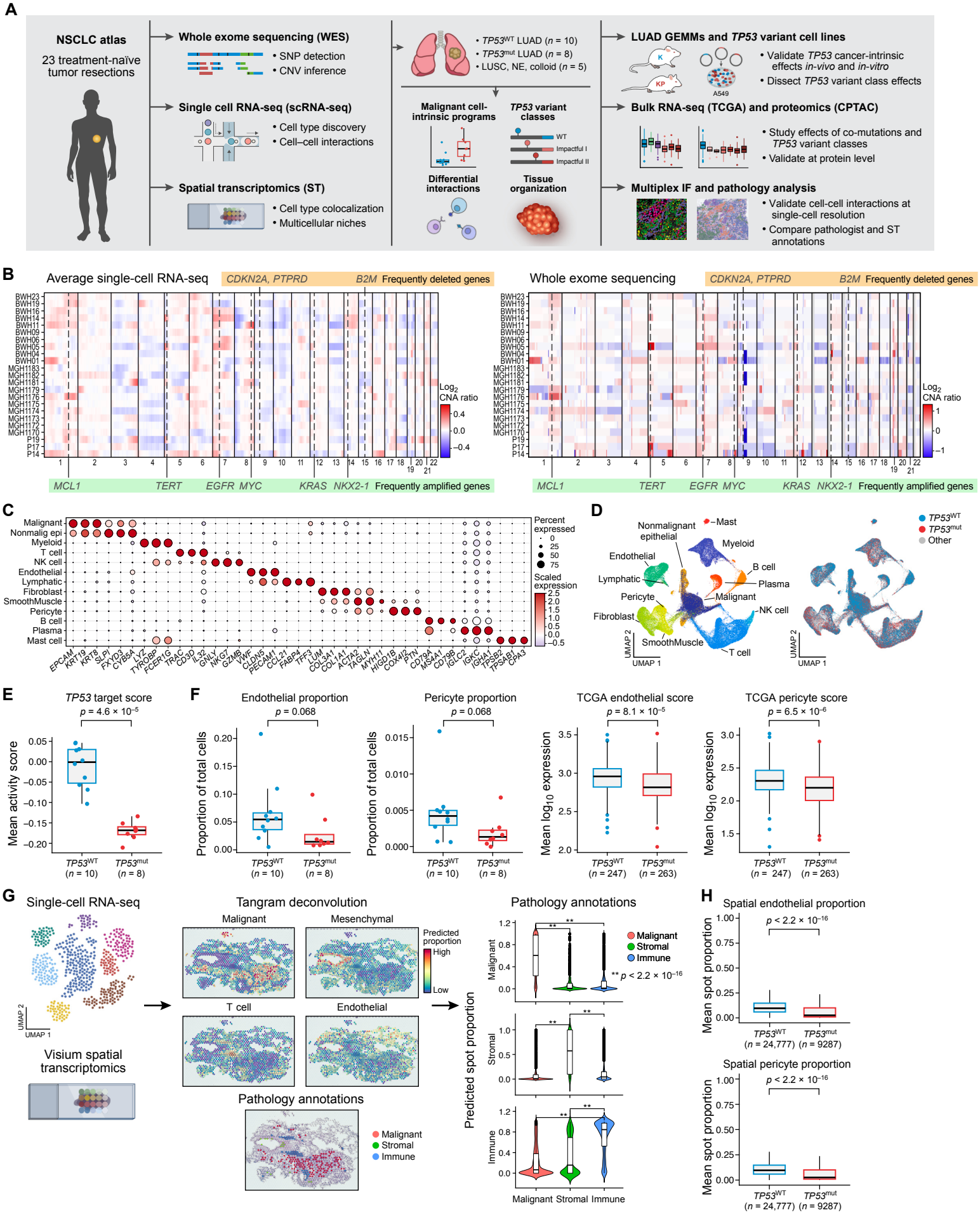
A
NSCLC atlas
23 treatment-naïve tumor resections
Whole exome sequencing (WES)
• SNP detection
• CNV inference
Single cell RNA-seq (scRNA-seq)
• Cell type discovery
• Cell–cell interactions
Spatial transcriptomics (ST)
• Cell type colocalization
• Multicellular niches
• TP53WT LUAD (n = 10)
• TP53mut LUAD (n = 8)
• LUSC, NE, colloid (n = 5)
Malignant cell-intrinsic programs
TP53 variant classes
WT
Impactful I
Impactful II
Differential interactions
Tissue organization
LUAD GEMMs and TP53 variant cell lines
• Validate TP53 cancer-intrinsic effects in-vivo and in-vitro
• Dissect TP53 variant class effects
A549
Bulk RNA-seq (TCGA) and proteomics (CPTAC)
• Study effects of co-mutations and TP53 variant classes
• Validate at protein level
Multiplex IF and pathology analysis
• Validate cell-cell interactions at single-cell resolution
• Compare pathologist and ST annotations
B
Average single-cell RNA-seq
CDKN2A, PTPRD
B2M Frequently deleted genes
Log2 CNA ratio
MCL1
TERT
EGFR MYC
KRAS NKX2-1 Frequently amplified genes
Whole exome sequencing
C
Malignant
Nonmalig epi
Myeloid
T cell
NK cell
Endothelial
Lymphatic
Fibroblast
SmoothMuscle
Pericyte
B cell
Plasma
Mast cell
Percent expressed
Scaled expression
D
Nonmalignant epithelial
Mast
Myeloid
Endothelial
B cell
Lymphatic
Plasma
Pericyte
Malignant
NK cell
Fibroblast
SmoothMuscle
T cell
UMAP 1
UMAP 2
TP53WT
TP53mut
Other
E
TP53 target score
p = 4.6 × 10−5
Mean activity score
TP53WT (n = 10)
TP53mut (n = 8)
F
Endothelial proportion
p = 0.068
Proportion of total cells
Pericyte proportion
p = 0.068
TCGA endothelial score
p = 8.1 × 10−5
Mean log10 expression
TP53WT (n = 247)
TP53mut (n = 263)
TCGA pericyte score
p = 6.5 × 10−6
G
Single-cell RNA-seq
Tangram deconvolution
Malignant
Mesenchymal
T cell
Endothelial
Predicted proportion
High
Low
Visium spatial transcriptomics
Pathology annotations
Malignant
Stromal
Immure
Predicted spot proportion
** p < 2.2 × 10−16
H
Spatial endothelial proportion
p < 2.2 × 10−16
Mean spot proportion
TP53WT (n = 24,777)
TP53mut (n = 9287)
Spatial pericyte proportion

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

A. Schematic of study design, including cohort selection (left), genomic assays profiled, sample stratification, downstream analyses, and validation platforms (right).

B. Mean $\log_2$ copy number alteration (CNA) ratio inferred using scRNA-seq malignant cells (left) and bulk WES (right) from tumor-matched samples. Common focal amplifications (bottom, green) and deletions (top, orange) in LUAD are highlighted.

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

D. Uniform Manifold Approximation and Projection (UMAP) visualization of 167,193 cells, integrated by tumor sample to minimize intertumoral heterogeneity and colored by annotated cell classes (left) and *TP53* mutational status (right).

E. Average *TP53* target gene expression score in malignant cells isolated from $TP53^{WT}$ (blue) and $TP53^{mut}$ (red) LUAD primary tumor samples.

F. Left: Proportion of endothelial cells and pericytes from scRNA-seq of $TP53^{WT}$ (blue) and $TP53^{mut}$ (red) LUAD tumors. Right: Mean $\log_{10}$ expression of highly specific endothelial and pericyte markers derived from scRNA-seq in bulk RNA-seq data from TCGA. *P*-values were calculated using a two-tailed multiple regression t-test.

G. Left: Integration of ST data with tumor-matched scRNA-seq data using Tangram allows for deconvolution of the proportion of different cell subsets. Middle: Spatial distribution of cell classes using Tangram (top), and a representative H&E stain annotated by a 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 ST spots from *TP53*^{WT} (blue) and *TP53*^{mut} (red) LUAD sections. Mann-Whitney-Wilcoxon test *p*-values were computed.

Figure 2

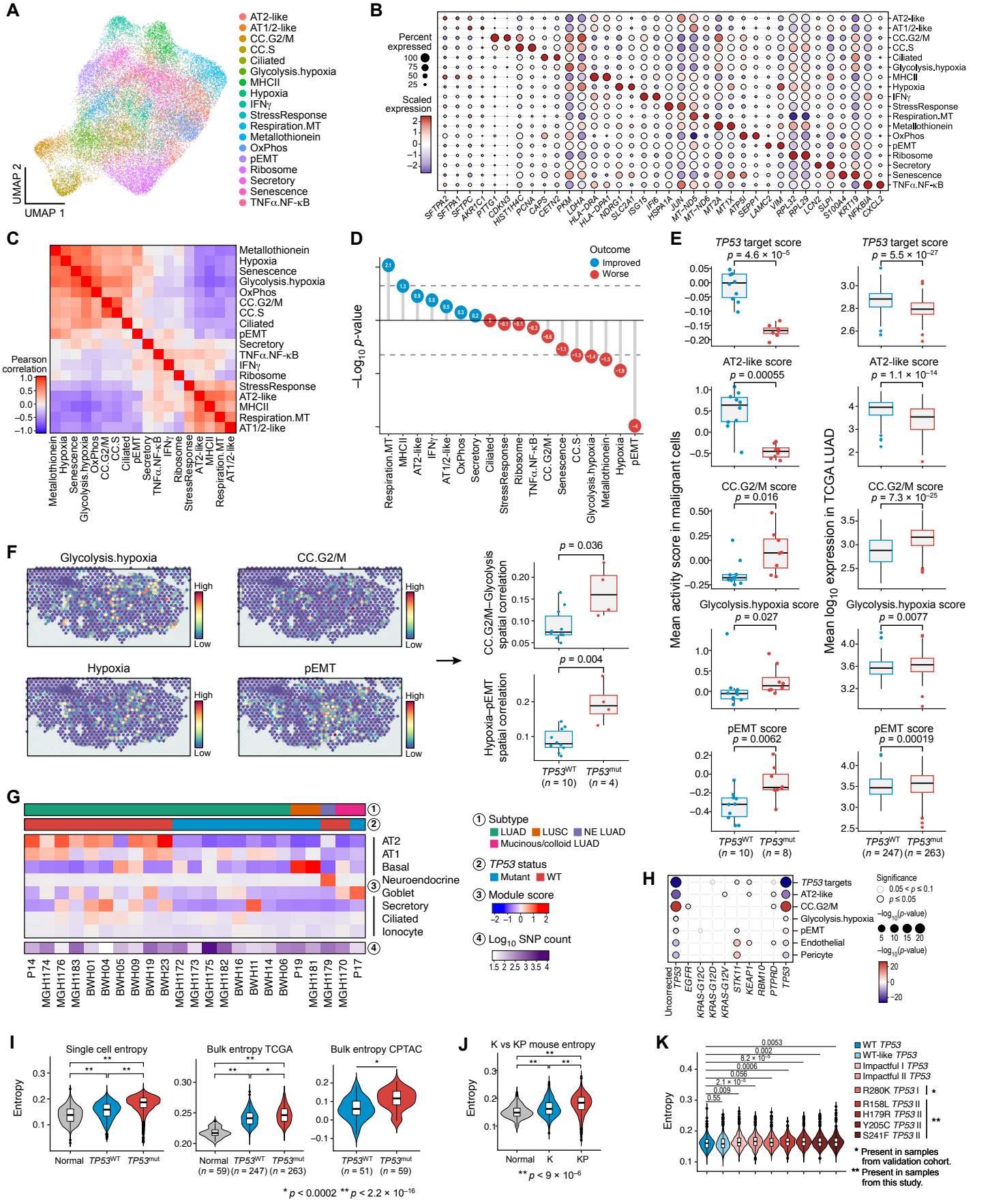
A
AT2-like
AT1/2-like
CC.G2/M
CC.S
Ciliated
Glycolysis.hypoxia
MHCII
Hypoxia
IFNγ
StressResponse
Respiration.MT
Metallothionein
OxPhos
pEMT
Ribosome
Secretory
Senescence
TNFα.NF-κB
UMAP 2
UMAP 1
B
Percent expressed
Scaled expression
C
Pearson correlation
D
Outcome
Improved
Worse
−Log10 p-value
E
TP53 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 = 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
TP53WT (n = 10)
TP53mut (n = 4)
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
TP53 targets
AT2-like
CC.G2/M
Glycolysis.hypoxia
pEMT
Endothelial
Pericyte
Significance
Uncorrected
I
Single cell entropy
Bulk entropy TCGA
Bulk entropy CPTAC
Entropy
Normal
TP53WT
TP53mut
* p < 0.0002 ** p < 2.2 × 10−16
J
K vs KP mouse entropy
Normal
K
KP
** p < 9 × 10−6
K
WT TP53
WT-like TP53
Impactful I TP53
Impactful II TP53
R280K TP53 I
R158L TP53 II
H179R TP53 II
Y205C TP53 II
S241F TP53 II
* Present in samples from validation cohort.
** Present in samples from this study.

Figure 2. Enrichment of cell cycle, hypoxia and pEMT programs accompanied by loss of AT2 identity in *TP53*^{mut} LUAD malignant cells.

A. UMAP of 33,377 malignant cells integrated across tumors and colored by annotated malignant expression program subsets.

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

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

D. Lollipop plot showing -$\log_{10}$ *p*-values from Cox proportional-hazards regression analysis linking malignant programs to TCGA patient disease outcome, corrected for *TP53* mutational status. Sign of the y-axis corresponds to improved (positive value) or worse (negative value) outcome.

E. Comparison of mean malignant program expression scores between *TP53*^{mut} and *TP53*^{WT} tumors in scRNA-seq (left) and in bulk RNA-seq data from TCGA (right). *P*-values were calculated using a Mann-Whitney-Wilcoxon test (scRNA-seq, left) or a two-tailed multiple regression t-test (TCGA, right).

F. Left: Representative example of the spatial distribution of four malignant programs' expression in a *TP53*^{mut} patient. Right: Bar plots comparing colocalization (spatial correlation) of pairs of malignant programs between *TP53*^{mut} and *TP53*^{WT} tumor sections. Mann-Whitney-Wilcoxon test *p*-values were computed.

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

H. Dot plot showing association of different mutations (columns) with mean $\log_{10}$ expression of malignant programs, endothelial, and pericyte scores (rows) in TCGA LUAD. Color and size of

dots correspond to -$\log_{10}$ *p*-value assessed by Mann-Whitney-Wilcoxon (uncorrected *TP53*, left) or a two-tailed multiple regression t-test (*EGFR*, *KRAS-G12C*, *KRAS-G12D*, *KRAS-G12V*, *STK11*, *KEAP1*, *RBM10*, *PTPRD*, *TP53*). Black outline indicates *p*-value ≤ 0.05, and gray outline indicates *p*-value of ≤ 0.1.

I. 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), across adjacent normal (grey), $TP53^{WT}$ (blue), and $TP53^{mut}$ (red) bulk transcriptomic LUAD tumor samples from TCGA (middle), and across $TP53^{WT}$ (blue), and $TP53^{mut}$ bulk proteomic tumor samples from CPTAC (right). *P*-values were calculated using a Mann-Whitney-Wilcoxon test.

J. Distribution of entropy scores across normal cells (grey), malignant $Kras^{mut}$/ $Trp53^{WT}$ cells (K; blue), and malignant $Kras^{mut}$/$Trp53^{mut}$ cells (KP; red) in K and KP mouse tumors.

K. Comparison of entropy scores between $TP53^{WT}$ and different $TP53^{mut}$ impact categories in A549 cells. Rightmost four violin plots show the entropy distributions of mutations also found in our cohort.

Figure 3

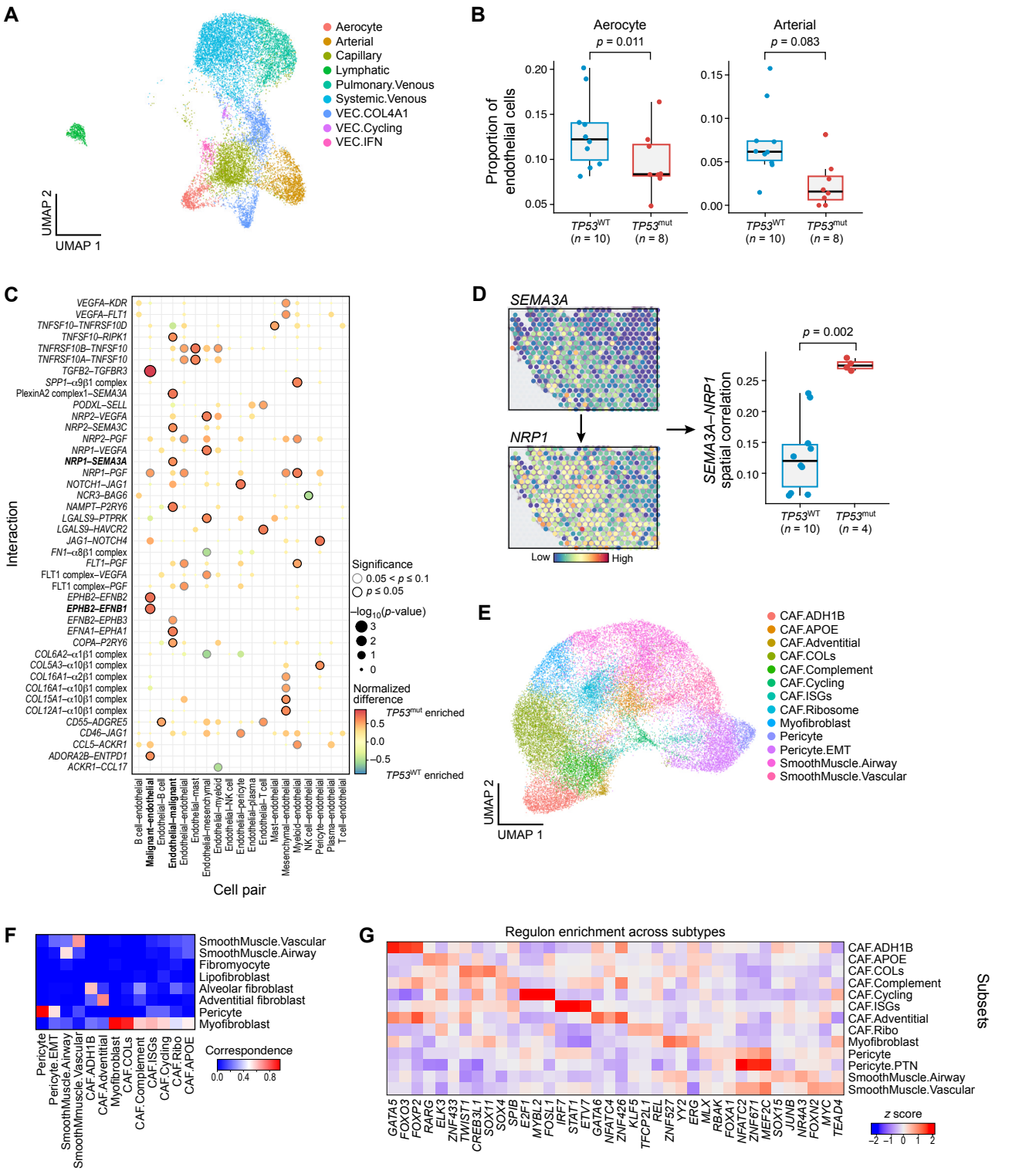
A
Aerocyte
Arterial
Capillary
Lymphatic
Pulmonary.Venous
Systemic.Venous
VEC.COL4A1
VEC.Cycling
VEC.IFN
UMAP 2
UMAP 1
B
Aerocyte
p = 0.011
Arterial
p = 0.083
Proportion of endothelial cells
TP53WT (n = 10)
TP53mut (n = 8)
C
Interaction
Cell pair
Significance
0.05 < p ≤ 0.1
p ≤ 0.05
−log10(p-value)
Normalized difference
TP53mut enriched
TP53WT enriched
D
SEMA3A
NRP1
Low
High
p = 0.002
SEMA3A–NRP1 spatial correlation
TP53WT (n = 10)
TP53mut (n = 4)
E
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
F
SmoothMuscle.Vascular
SmoothMuscle.Airway
Fibromyocyte
Lipofibroblast
Alveolar fibroblast
Adventitial fibroblast
Pericyte
Myofibroblast
Correspondence
G
Regulon enrichment across subtypes
Subsets
z score
H
Pericyte
p = 0.0044
CAF.ADH1B
p = 0.055
Proportion of mesenchymal cells
TP53WT (n = 10)
TP53mut (n = 8)
I
TGFB2
TGFBR2
Low
High
p = 0.002
TGFB2–TGFBR2 spatial correlation
TP53WT (n = 10)
TP53mut (n = 4)

Figure 3. Spatially-resolved interactions linked to vascular depletion and stromal remodeling in *TP53*^{mut} LUAD.

A. UMAP of 15,320 endothelial cells, integrated across tumors and colored by annotated subset.

B. Comparison of aerocyte (left) and arterial (right) cell subset proportions in *TP53*^{mut} and *TP53*^{WT}. Mann-Whitney-Wilcoxon test was used for computing *p*-values.

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

D. Left: Spatial expression of *SEMA3A* ligand and *NRP1* receptor in a representative *TP53*^{mut} ST sample. Right: Box plot comparing spatial correlation of *SEMA3A* and *NRP1* expression between *TP53*^{WT} and *TP53*^{mut} ST tumor samples. Mann-Whitney-Wilcoxon test *p*-value was computed.

E. UMAP of 31,868 mesenchymal cells integrated across tumors and colored by annotated subset.

F. Confusion matrix of correspondence between our *de novo* mesenchymal subset annotations and predicted identities using a classifier trained on annotated scRNA-seq data from normal lung mesenchymal cells.

G. Z-scores of the most specific SCENIC regulon activity for each mesenchymal subset.

H. Comparison of mesenchymal subset cell proportions in *TP53*^{mut} and *TP53*^{WT}. Mann-Whitney-Wilcoxon test *p*-values were computed.

I. Left: Spatial expression of *TGFB2* ligand and *TGFBR2* receptor in a representative *TP53*^{mut} ST sample. Right: Box plot comparing spatial correlation of *TGFB2* and *TGFBR2* expression between *TP53*^{WT} and *TP53*^{mut} ST tumor samples. Mann-Whitney-Wilcoxon test *p*-values were computed.

Figure 4

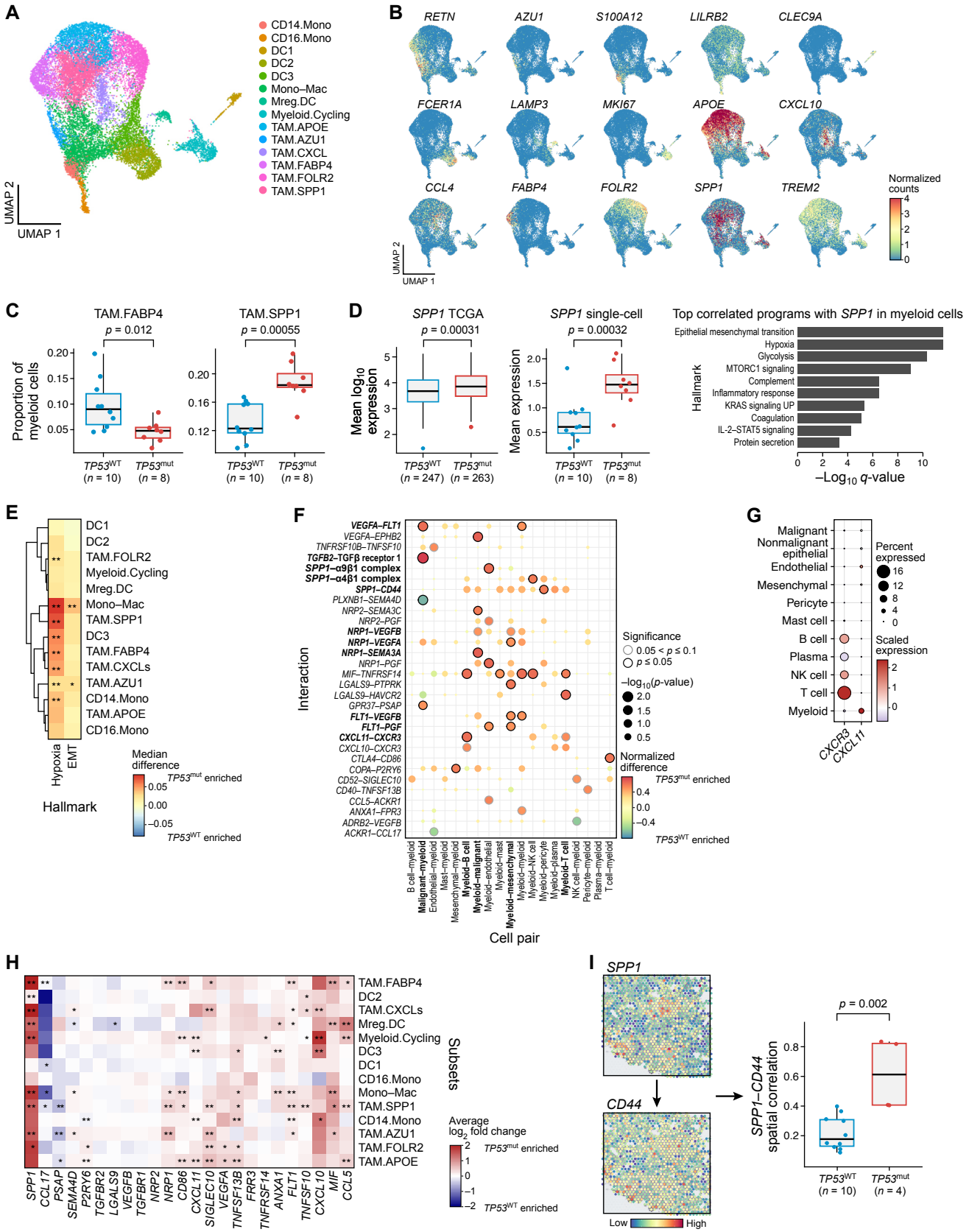
A
CD14.Mono
CD16.Mono
DC1
DC2
DC3
Mono–Mac
Mreg.DC
Myeloid.Cycling
TAM.APOE
TAM.AZU1
TAM.CXCL
TAM.FABP4
TAM.FOLR2
TAM.SPP1
UMAP 2
UMAP 1
B
RETN
AZU1
S100A12
LILRB2
CLEC9A
FCER1A
LAMP3
MKI67
APOE
CXCL10
CCL4
FABP4
FOLR2
SPP1
TREM2
Normalized counts
C
TAM.FABP4
p = 0.012
TAM.SPP1
p = 0.00055
Proportion 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 ≤ 0.1
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
CXCR3
CXCL11
Percent expressed
Scaled expression
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 4. *SPP1* and other immunomodulatory factors are enriched in *TP53*^{mut} LUAD myeloid compartment.

A. UMAP of 18,086 myeloid cells integrated across tumors and colored by annotated subset.

B. Feature plot showing the expression of 15 representative myeloid subset markers.

C. Comparison of myeloid subset cell proportions in *TP53*^{mut} and *TP53*^{WT} scRNA-seq tumors. 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} *vs*. *TP53*^{WT} primary tumor samples. Right: Horizontal bar plot of negative $\log_{10}$ *q*-values from gene set enrichment analysis of the 100 most positively correlated genes with *SPP1* in myeloid cells. *P*-values were calculated using a two-tailed multiple regression t-test (TCGA, left) or a Mann-Whitney-Wilcoxon test (scRNA-seq, middle).

E. Differential enrichment of hallmark hypoxia and EMT module score for each myeloid subset, comparing pseudobulk subset profiles for *TP53*^{mut} *vs*. *TP53*^{WT} tumors. Mann-Whitney-Wilcoxon test *p*-values were calculated.

F. Dot plot of differential ligand-receptor (rows) interactions between myeloid cells and other cell classes (columns). Red indicates enrichment in *TP53*^{mut} and blue enrichment in *TP53*^{WT} tumor 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 broad cell class annotations.

H. Differential expression analysis of ligand and receptor genes from the interactions displayed in **Figure 5F,** comparing pseudobulk myeloid subset profiles for *TP53*^{mut} *vs*. *TP53*^{WT} tumors. Mann-Whitney-Wilcoxon test *p*-values were calculated. Red indicates enrichment in *TP53*^{mut} and blue enrichment in *TP53*^{WT} tumor samples.

I. Left: Spatial expression of *SPP1* ligand and *CD44* receptor in a representative *TP53*^{mut} ST sample. Right: Box plot comparing spatial correlation of *SPP1* and *CD44* expression between *TP53*^{WT} and *TP53*^{mut} ST tumor samples. Mann-Whitney-Wilcoxon test *p*-values were computed.

Figure 5

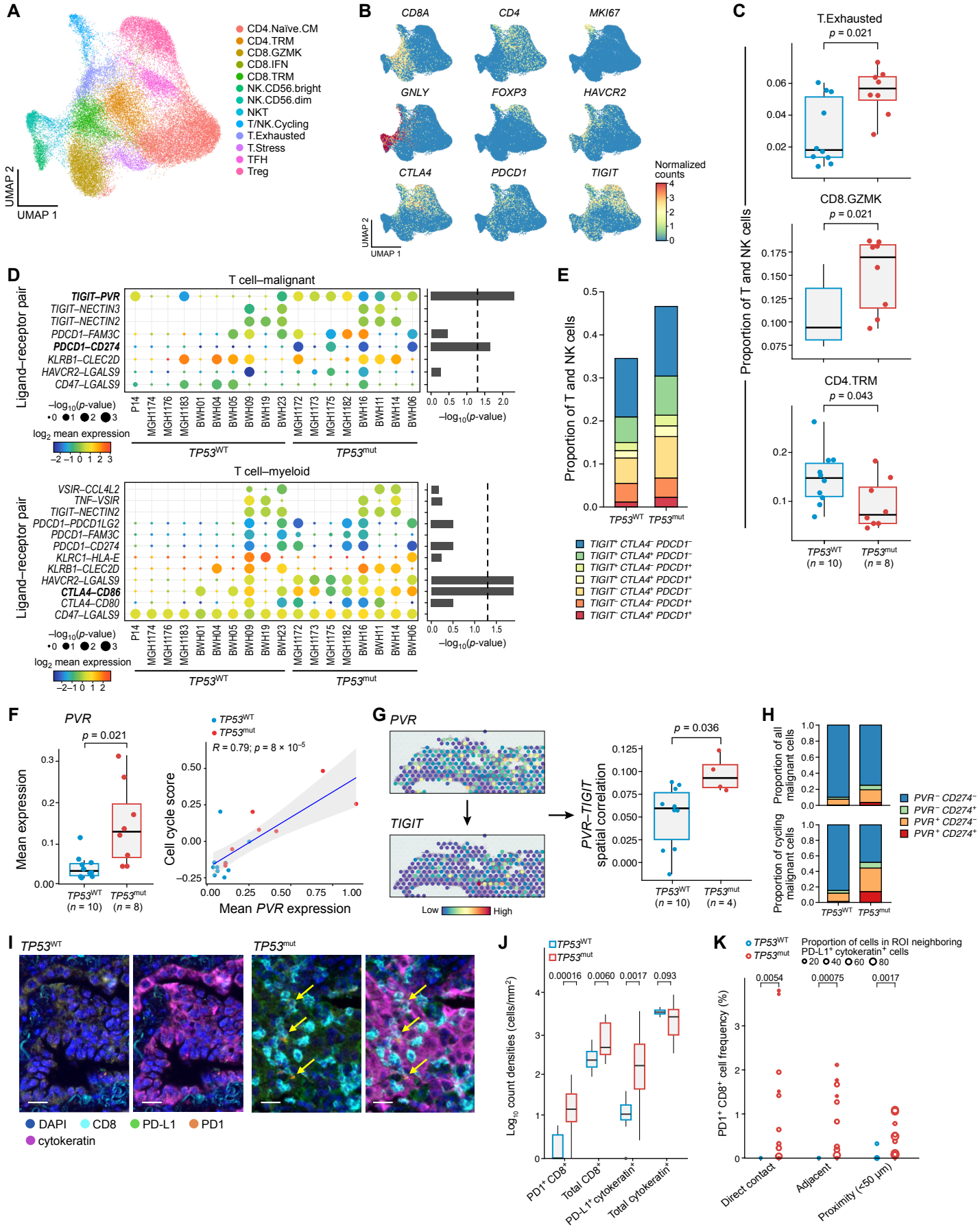
A
CD4.Naïve.CM
CD4.TRM
CD8.GZMK
CD8.IFN
CD8.TRM
NK.CD56.bright
NK.CD56.dim
NKT
T/NK.Cycling
T.Exhausted
T.Stress
TFH
Treg
UMAP 2
UMAP 1
B
CD8A
CD4
MKI67
GNLY
FOXP3
HAVCR2
CTLA4
PDCD1
TIGIT
Normalized counts
C
T.Exhausted
p = 0.021
CD8.GZMK
p = 0.021
CD4.TRM
p = 0.043
Proportion of T and NK cells
TP53WT (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 expression
TP53WT
TP53mut
T cell–myeloid
VSIR–CCL4L2
TNF–VSIR
TIGIT–NECTIN2
PDCD1–PDCD1LG2
PDCD1–FAM3C
PDCD1–CD274
KLRC1–HLA-E
KLRB1–CLEC2D
HAVCR2–LGALS9
CTLA4–CD86
CTLA4–CD80
CD47–LGALS9
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)
H
Proportion of all malignant cells
Proportion of cycling malignant cells
PVR− CD274−
PVR− CD274+
PVR+ CD274−
PVR+ CD274+
I
TP53WT
TP53mut
DAPI
CD8
PD-L1
PD1
cytokeratin
J
0.00016
0.0060
0.0017
0.093
Log10 count densities (cells/mm2)
PD1+ CD8+
Total CD8+
PD-L1+ cytokeratin+
Total cytokeratin+
K
Proportion of cells in ROI neighboring PD-L1+ cytokeratin+ cells
20
40
60
80
0.0054
0.00075
0.0017
PD1+ CD8+ cell frequency (%)
Direct contact
Adjacent
Proximity (<50 μm)

Figure 5. 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 tumors and colored by annotated subset.

B. Feature plot showing the expression of 9 representative T and NK cell subset markers.

C. Comparison of T.Exhausted, CD8.GZMK, and CD4.TRM T cell subset proportions relative to all T and NK cells in *TP53*^{mut} and *TP53*^{WT} scRNA-seq tumors. Mann-Whitney-Wilcoxon test *p*-values were computed.

D. Log_2 mean gene expression and predicted significance of ICI-targetable ligand-receptor interactions between T cells and malignant cells (top), and T cells and myeloid (bottom) for each tumor. Horizontal bar plots (right) show negative log_{10} *p*-value (Fisher's exact test) of the enrichment in interactions between *TP53*^{mut} and *TP53*^{WT} tumors.

E. Stacked bar plots depicting proportion of T and NK cells expressing different combinations of immune checkpoint molecules *TIGIT*, *CTLA4*, and *PDCD1* in *TP53*^{WT} (left) and *TP53*^{mut} (right) tumors.

F. Left: Boxplot of mean *PVR* expression in *TP53*^{WT} (blue) and *TP53*^{mut} (red) tumor malignant cells. Right: Correlation of cell cycle score *vs*. mean *PVR* expression in malignant cells for each patient, where blue represents *TP53*^{WT} and red represents *TP53*^{mut} tumors.

G. Left: Spatial expression of *PVR* and *TIGIT* in a representative *TP53*^{mut} ST sample. Right: Box plot comparing spatial correlation of *PVR* and *TIGIT* expression between *TP53*^{WT} and *TP53*^{mut} tumor ST samples (right) in the tumor periphery (spots containing ≤50% malignant cells). Mann-Whitney-Wilcoxon test *p*-values were computed.

H. Stacked bar plots depicting proportion of malignant cells expressing different combination of *PVR* and *CD274* amongst all malignant cells (top) and amongst all cycling malignant cells (bottom) in *TP53*^{WT} (left) and *TP53*^{mut} (right) tumors.

I. Representative mIF images of *TP53*^{WT} (left two; BWH01) and *TP53*^{mut} (right two; BWH06) tumors with and without cytokeratin (pink). Yellow arrows correspond to PD1^{+}CD8^{+} cells. Scale bar, 20 µm.

J. Box plots showing the $\log_{10}$ count density (cells/mm^2) of PD1^{+}CD8^{+}, total CD8^{+}, PD-L1^{+}cytokeratin^{+}, and total cytokeratin^{+} cells across regions of interest (ROIs) in *TP53*^{WT} (BWH01, BWH04, blue) and *TP53*^{mut} (BWH06, BWH11, red) tumor samples used for mIF. Mann-Whitney-Wilcoxon test *p*-values were computed.

K. Grouped scatter plot showing the proportion of PD1^{+}CD8^{+} cells out of all cells in proximity to PD-L1^{+}cytokeratin^{+}. Each circle represents an ROI from *TP53*^{WT} (BWH01, BWH04, blue) *vs*. *TP53*^{mut} (BWH06, BWH11, red) mIF tumor stains, where the size of the circle represents the proportion of PD-L1^{+}cytokeratin^{+}-neighboring cells, out of all cells within the ROI. Permutation Test *p*-values were computed.

Figure 6

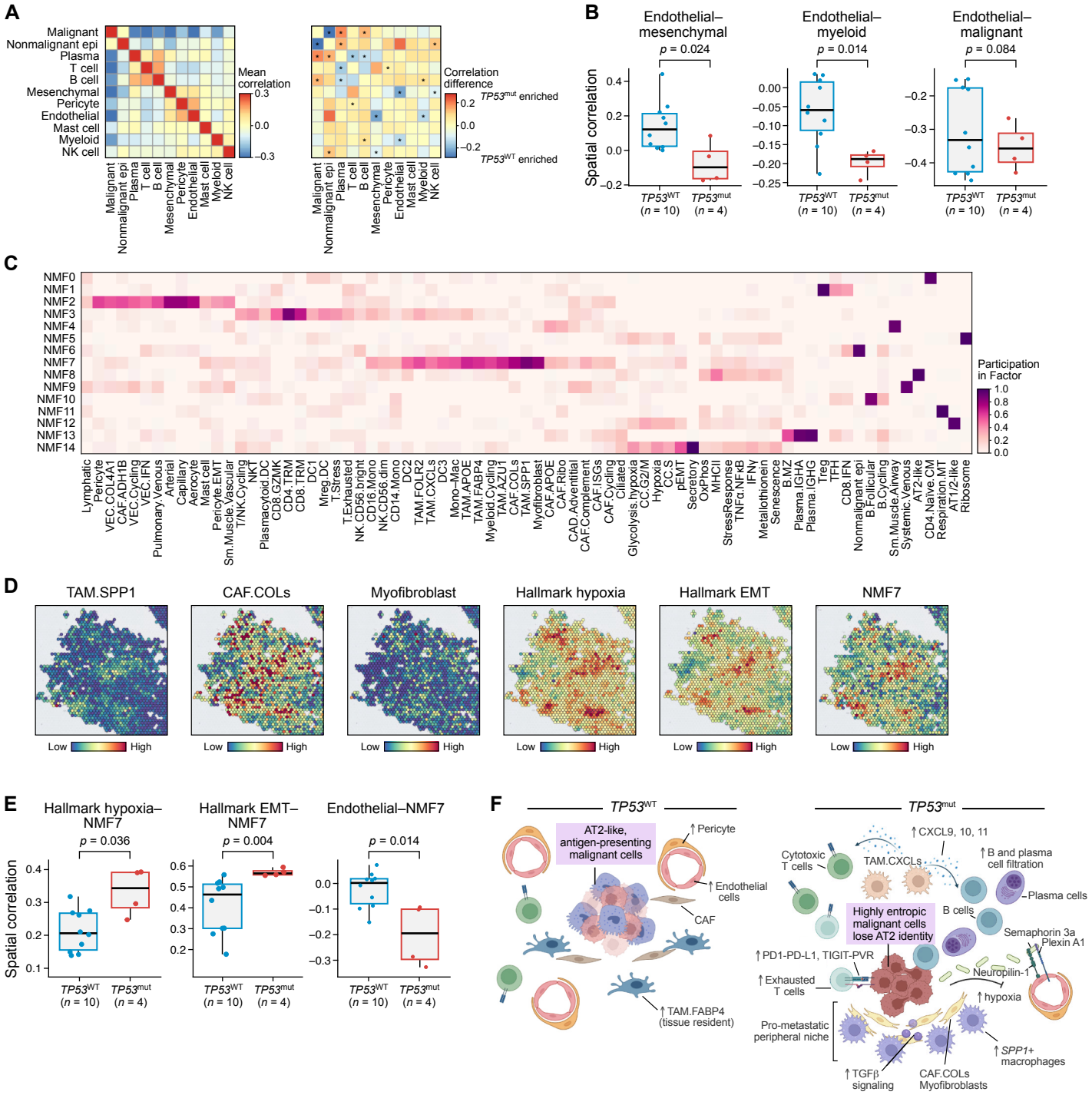
A
Malignant
Nonmalignant epi
Plasma
T cell
B cell
Mesenchymal
Pericyte
Endothelial
Mast cell
Myeloid
NK cell
Mean correlation
Correlation difference
TP53mut enriched
TP53WT enriched
B
Endothelial–mesenchymal
p = 0.024
Endothelial–myeloid
p = 0.014
Endothelial–malignant
p = 0.084
Spatial correlation
TP53WT (n = 10)
TP53mut (n = 4)
C
NMF0
NMF1
NMF2
NMF3
NMF4
NMF5
NMF6
NMF7
NMF8
NMF9
NMF10
NMF11
NMF12
NMF13
NMF14
Participation in Factor
Lymphatic
Pericyte
VEC.COL4A1
CAF.ADH1B
VEC.Cycling
VEC.IFN
Pulmonary.Venous
Arterial
Capillary
Aerocyte
Mast cell
Pericyte.EMT
Sm.Muscle.Vascular
T/NK.Cycling
NKT
Plasmacytoid.DC
CD8.GZMK
CD4.TRM
CD8.TRM
DC1
Mreg.DC
T.Stress
T.Exhausted
NK.CD56.bright
CD16.Mono
NK.CD56.dim
CD14.Mono
DC2
TAM.FOLR2
TAM.CXCLs
DC3
Mono–Mac
TAM.APOE
TAM.FABP4
Myeloid.Cycling
TAM.AZU1
CAF.COLs
TAM.SPP1
Myofibroblast
CAF.APOE
CAF.Ribo
CAD.Adventitial
CAF.Complement
CAF.ISGs
CAF.Cycling
Ciliated
Glycolysis.hypoxia
CC.G2/M
Hypoxia
CC.S
pEMT
Secretory
OxPhos
MHCII
StressResponse
TNFα.NFκB
IFNγ
Metallothionein
Senescence
B.MZ
Plasma.IGHA
Plasma.IGHG
Treg
TFH
CD8.IFN
Nonmalignant.epi
B.Follicular
B.Cycling
Sm.Muscle.Airway
Systemic.Venous
AT2-like
CD4.Naïve.CM
Respiration.MT
AT1/2-like
Ribosome
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
TP53WT (n = 10)
TP53mut (n = 4)
F
TP53WT
TP53mut
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
Semaphorin 3a
Plexin A1
PD1-PD-L1, TIGIT-PVR
Neuropilin-1
Exhausted T cells
hypoxia
Pro-metastatic peripheral niche
SPP1+ macrophages
TGFβ signaling
CAF.COLs
Myofibroblasts

Figure 6. Spatial multicellular CAF-TAM niche enriched in hypoxia and EMT activity in *TP53*^{mut} LUAD.

A. Overall spatial correlation of broad cell classes across ST spots (left) and differential spatial correlation between *TP53*^{mut} and *TP53*^{WT} tumor samples (right). Red indicates either high colocalization (left) or colocalization enrichment in *TP53*^{mut} (right), while blue represents either low colocalization (left) or colocalization enrichment in *TP53*^{WT} (right) tumor samples between pairs of cell subsets.

B. Box plots comparing spatial correlation of endothelial cells with mesenchymal (left), myeloid (middle), and malignant (right) cells across spots in *TP53*^{WT} *vs*. *TP53*^{mut} ST tumor 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 cell subsets comprising NMF7 (left) and of expression programs spatially correlated with NMF7 (right) in a representative *TP53*^{mut} tumor ST sample (patient BWH14).

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} *vs*. *TP53*^{mut} ST samples in the tumor periphery (spots with ≤50% malignant cells).

F. Graphical summary of the differences between the *TP53*^{WT} (left) and *TP53*^{mut} (right) LUAD TME presented in this study.

Figure S1

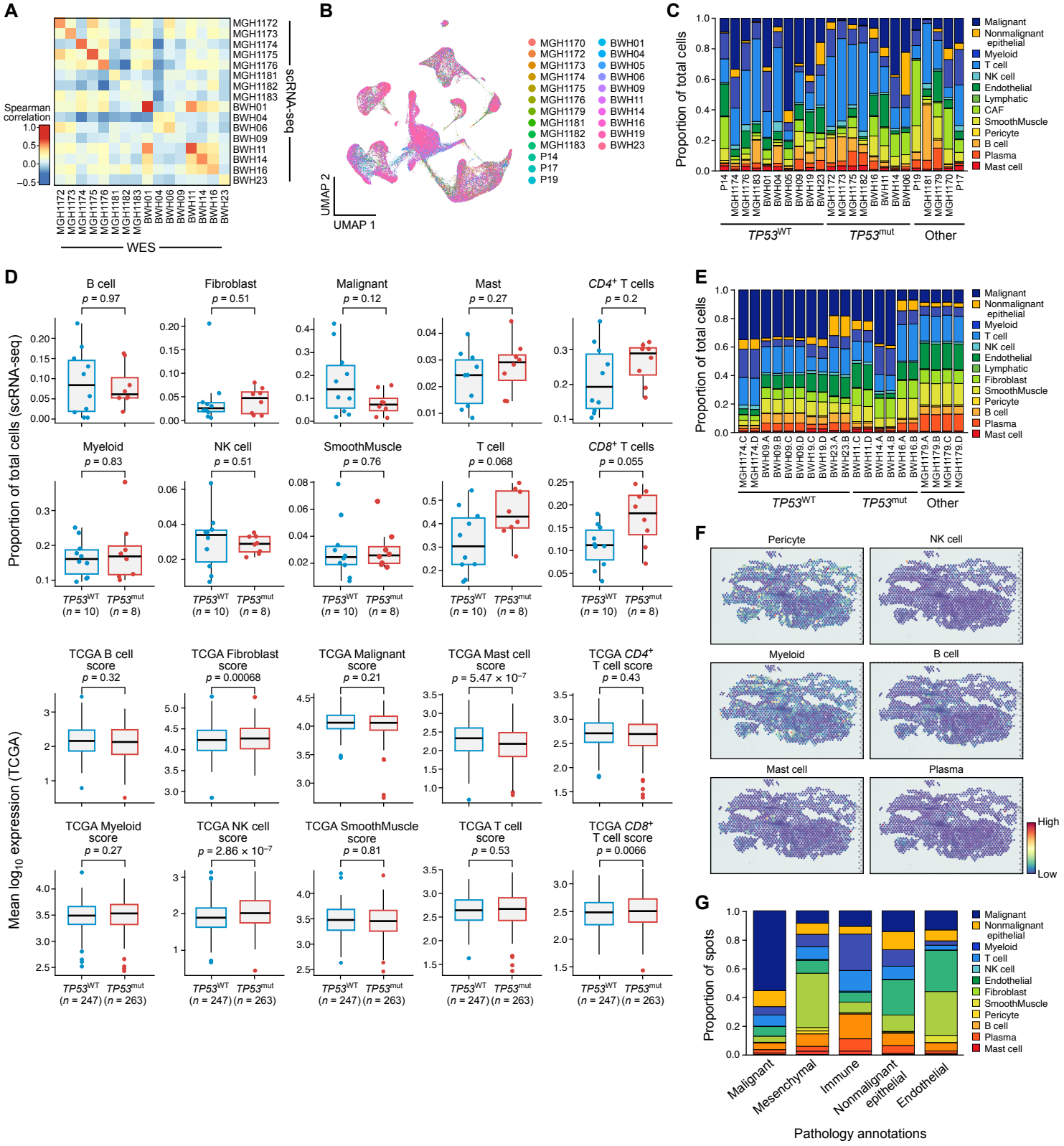
A
Spearman correlation
scRNA-seq
WES
B
UMAP 1
UMAP 2
C
Proportion of total cells
Malignant
Nonmalignant epithelial
Myeloid
T cell
NK cell
Endothelial
Lymphatic
CAF
SmoothMuscle
Pericyte
B cell
Plasma
Mast cell
TP53WT
TP53mut
Other
D
B cell p = 0.97
Fibroblast p = 0.51
Malignant p = 0.12
Mast p = 0.27
CD4+ T cells p = 0.2
Myeloid p = 0.83
NK cell p = 0.51
SmoothMuscle p = 0.76
T cell p = 0.068
CD8+ T cells p = 0.055
Proportion of total cells (scRNA-seq)
TCGA B cell score p = 0.32
TCGA Fibroblast score p = 0.00068
TCGA Malignant score p = 0.21
TCGA Mast cell score p = 5.47 × 10−7
TCGA CD4+ T cell score p = 0.43
TCGA Myeloid score p = 0.27
TCGA NK cell score p = 2.86 × 10−7
TCGA SmoothMuscle score p = 0.81
TCGA T cell score p = 0.53
TCGA CD8+ T cell score p = 0.0066
Mean log10 expression (TCGA)
TP53WT (n = 10)
TP53mut (n = 8)
TP53WT (n = 247)
TP53mut (n = 263)
E
Proportion of total cells
F
Pericyte
NK cell
Myeloid
B cell
Mast cell
Plasma
High
Low
G
Proportion of spots
Malignant
Mesenchymal
Immune
Nonmalignant epithelial
Endothelial
Pathology annotations

Figure S1. Integration and analysis of WES, scRNA-seq, and spatial transcriptomics data across tumors.

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. Low correspondence in two tumor samples (MGH1183 and BWH09) likely resulted from small number of malignant cells (MGH1183) and intratumoral heterogeneity (BWH09).

B. UMAP of 167,193 cells, colored by patient identity shows integration between tumor samples.

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

D. Top: Proportion of broad cell classes (not shown in **Figure 1F**) out of all cells found in the scRNA-seq data from *TP53*^{WT} (blue) and *TP53*^{mut} (red) LUAD tumors. 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 tumor samples from TCGA. *P*-values were calculated using a Mann-Whitney-Wilcoxon test (scRNA-seq, top) or a two-tailed multiple regression t-test (TCGA, bottom).

E. Stacked bar plot depicting mean cell composition across ST tumor samples, predicted using Tangram mapping.

F. Spatial feature plots of broad cell classes not shown in **Figure 1G**, predicted by Tangram mapping.

G. Mean proportion of Tangram-predicted cell classes (y-axis) across different manual spot annotations independently performed by a pathologist (NRM) without knowledge of sample and spot identities.

Figure S2

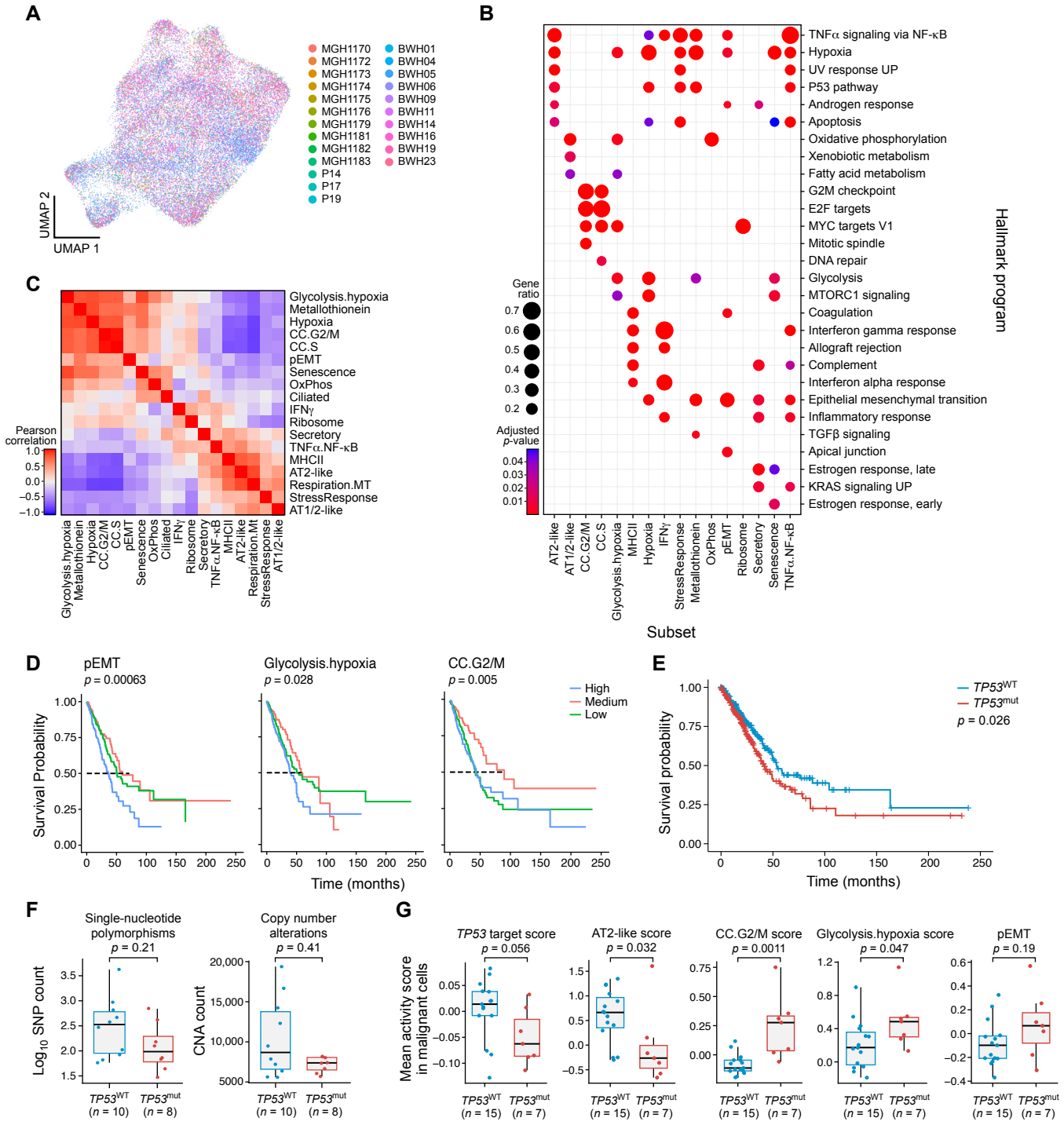
A
MGH1170
MGH1172
MGH1173
MGH1174
MGH1175
MGH1176
MGH1179
MGH1181
MGH1182
MGH1183
P14
P17
P19
BWH01
BWH04
BWH05
BWH06
BWH09
BWH11
BWH14
BWH16
BWH19
BWH23
UMAP 2
UMAP 1
B
TNFα signaling via NF-κB
Hypoxia
UV response UP
P53 pathway
Androgen response
Apoptosis
Oxidative phosphorylation
Xenobiotic metabolism
Fatty acid metabolism
G2M checkpoint
E2F targets
MYC targets V1
Mitotic spindle
DNA repair
Glycolysis
MTORC1 signaling
Coagulation
Interferon gamma response
Allograft rejection
Complement
Interferon alpha response
Epithelial mesenchymal transition
Inflammatory response
TGFβ signaling
Apical junction
Estrogen response, late
KRAS signaling UP
Estrogen response, early
Hallmark program
Gene ratio
Adjusted p-value
Subset
C
Pearson correlation
Glycolysis.hypoxia
Metallothionein
Hypoxia
CC.G2/M
CC.S
pEMT
Senescence
OxPhos
Ciliated
IFNγ
Ribosome
Secretory
TNFα.NF-κB
MHCII
AT2-like
Respiration.MT
StressResponse
AT1/2-like
D
pEMT
p = 0.00063
Glycolysis.hypoxia
p = 0.028
CC.G2/M
p = 0.005
High
Medium
Low
Survival Probability
Time (months)
E
TP53WT
TP53mut
p = 0.026
Survival probability
Time (months)
F
Single-nucleotide polymorphisms
p = 0.21
Log10 SNP count
Copy number alterations
p = 0.41
CNA count
TP53WT (n = 10)
TP53mut (n = 8)
G
TP53 target score
p = 0.056
AT2-like score
p = 0.032
CC.G2/M score
p = 0.0011
Glycolysis.hypoxia score
p = 0.047
pEMT
p = 0.19
Mean activity score in malignant cells
TP53WT (n = 15)
TP53mut (n = 7)

Figure S2. Characterization of the malignant compartment in NSCLC.

A. UMAP of 33,377 malignant cells, integrated across tumors 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 tumors, analogous to **Figure 2C** but performed across patients instead of across cells.

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

E. Kaplan Meier survival curves showing association of *TP53* mutation status with survival in bulk TCGA LUAD. *P*-values were calculated using a log-rank test.

F. Comparison of $\log_{10}$ SNP and CNA abundance between $TP53^{mut}$ *vs.* $TP53^{WT}$ tumors shows no significant difference in our cohort.

G. Comparison of malignant program mean expression scores between $TP53^{mut}$ and $TP53^{WT}$ tumors in scRNA-seq validation cohort. *P*-values were calculated using a Mann-Whitney-Wilcoxon test.

Figure S3

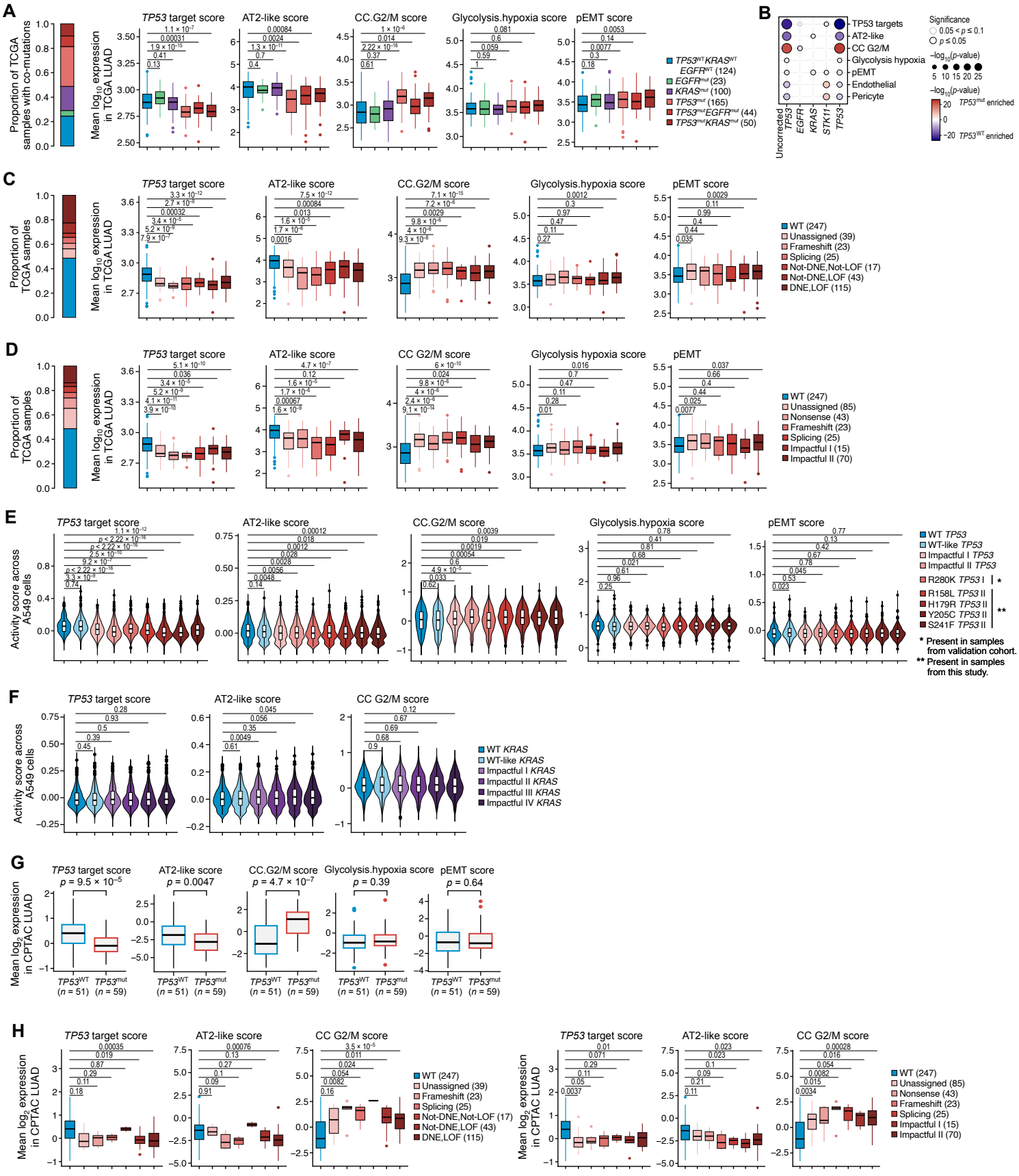
A
Proportion of TCGA samples with co-mutations
Mean log10 expression in TCGA LUAD
TP53 target score
AT2-like score
CC.G2/M score
Glycolysis.hypoxia score
pEMT score
TP53WT KRASWT EGFRWT (124)
EGFRmut (23)
KRASmut (100)
TP53mut (165)
TP53mut EGFRmut (44)
TP53mut KRASmut (50)
B
TP53 targets
AT2-like
CC G2/M
Glycolysis hypoxia
pEMT
Endothelial
Pericyte
Uncorrected
TP53
EGFR
KRAS
STK11
TP53
Significance
0.05 < p ≤ 0.1
p ≤ 0.05
−log10(p-value)
TP53mut enriched
TP53WT enriched
C
Proportion of TCGA samples
WT (247)
Unassigned (39)
Frameshift (23)
Splicing (25)
Not-DNE,Not-LOF (17)
Not-DNE,LOF (43)
DNE,LOF (115)
D
WT (247)
Unassigned (85)
Nonsense (43)
Frameshift (23)
Splicing (25)
Impactful I (15)
Impactful II (70)
E
Activity score across A549 cells
WT TP53
WT-like TP53
Impactful I TP53
Impactful II TP53
R280K TP53 I
R158L TP53 II
H179R TP53 II
Y205C TP53 II
S241F TP53 II
* Present in samples from validation cohort.
** Present in samples from this study.
F
WT KRAS
WT-like KRAS
Impactful I KRAS
Impactful II KRAS
Impactful III KRAS
Impactful IV KRAS
G
Mean log2 expression in CPTAC LUAD
TP53 target score p = 9.5 × 10−5
AT2-like score p = 0.0047
CC.G2/M score p = 4.7 × 10−7
Glycolysis.hypoxia score p = 0.39
pEMT score p = 0.64
TP53WT (n = 51)
TP53mut (n = 59)
H

Figure S3. Consistent changes in malignant program expression and cellular entropy in *TP53*^{mut} LUAD across co-mutations, variant classes, cell lines, mouse models, bulk transcriptomic, and proteomic data.

A. Left: Stacked bar plot of proportion of tumor samples with different co-mutations in TCGA. Right: Comparison of malignant program mean expression between *TP53*^{WT} and different single- or co-mutations in *EGFR*, *KRAS*, and *TP53* in bulk RNA-seq data from TCGA tumor samples.

B. Dot plot showing association of common LUAD mutations (columns) with mean $\log_{10}$ expression of malignant programs, endothelial, and pericyte scores (rows) in TCGA LUAD. Color and size of dots correspond to -log10 *p*-value assessed by Mann-Whitney-Wilcoxon (uncorrected *TP53*, left) or a two-tailed multiple regression t-test (*EGFR*, *KRAS, STK11*, *TP53*). Black outline indicates *p*-value ≤ 0.05, and gray outline indicates *p*-value of ≤ 0.1.

C. Left: Stacked bar plot of proportion of tumor samples with different categories of dominant-negative/loss of function *TP53*^{mut} effect in TCGA. Right: Comparison of malignant program mean expression between *TP53*^{WT} and different functional *TP53*^{mut} categories in TCGA tumor samples. *P*-values were calculated using a Mann-Whitney-Wilcoxon test.

D. Left: Stacked bar plot of proportion of tumor samples with different impact categories in TCGA. Right: Comparison of malignant program mean $\log_{10}$ expression between *TP53*^{WT} and different functional *TP53*^{mut} impact categories in TCGA tumor samples. *P*-values were calculated using a Mann-Whitney-Wilcoxon test.

E. Comparison of malignant program mean expression between *TP53*^{WT} and different *TP53*^{mut} impact categories[59] in A549 cells. Rightmost four violin plots in each panel show the effect of mutations also found in tumor samples from our cohort.

F. Comparison of malignant program mean expression scores between *KRAS*^{WT} and different *KRAS*^{mut} impact categories in A549 cells.

G. Comparison of malignant program mean expression between *TP53*^{mut} and *TP53*^{WT} tumors using bulk proteomic data from CPTAC. *P*-values were calculated using a Mann-Whitney-Wilcoxon test.

H. Comparison of malignant program mean $\log_2$ expression between *TP53*^{WT} and different dominant-negative/loss of function (left) and impact categories[59] (right) of *TP53*^{mut} tumor samples in bulk proteomic data from CPTAC. *P*-values were calculated using a Mann-Whitney-Wilcoxon test.

Figure S4

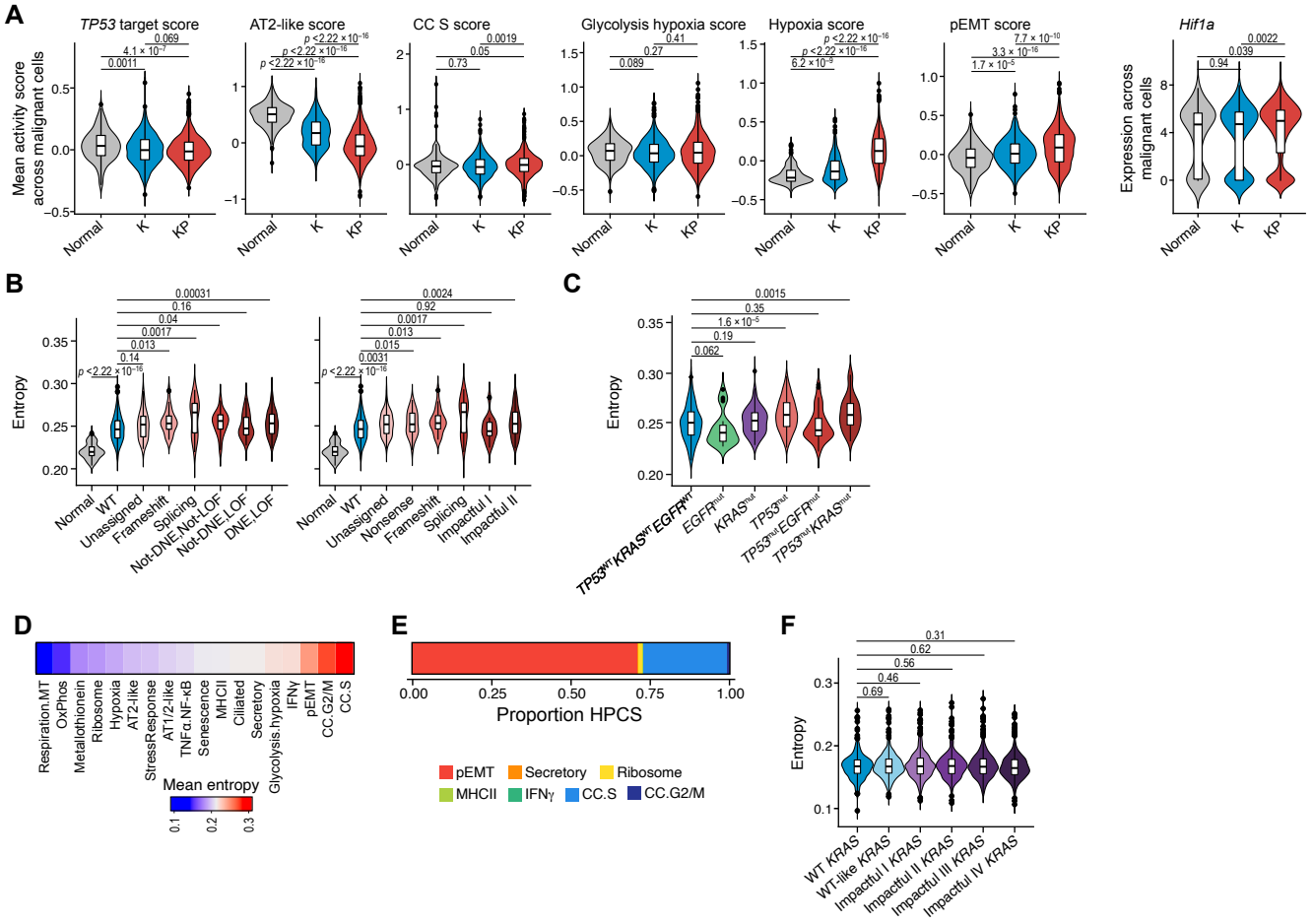
A
TP53 target score
AT2-like score
CC S score
Glycolysis hypoxia score
Hypoxia score
pEMT score
Hif1a
Mean activity score across malignant cells
Expression across malignant cells
Normal
K
KP
B
Entropy
Normal
WT
Unassigned
Frameshift
Splicing
Not-DNE,Not-LOF
Not-DNE,LOF
DNE,LOF
Nonsense
Impactful I
Impactful II
C
Entropy
TP53WT KRASWT EGFRWT
EGFRmut
KRASmut
TP53mut
TP53mut EGFRmut
TP53mut KRASmut
D
Respiration.MT
OxPhos
Metallothionein
Ribosome
Hypoxia
AT2-like
StressResponse
AT1/2-like
TNFα.NF-κB
Senescence
MHCII
Ciliated
Secretory
Glycolysis.hypoxia
IFNγ
pEMT
CC.G2/M
CC.S
Mean entropy
E
Proportion HPCS
pEMT
Secretory
Ribosome
MHCII
IFNγ
CC.S
CC.G2/M
F
Entropy
WT KRAS
WT-like KRAS
Impactful I KRAS
Impactful II KRAS
Impactful III KRAS
Impactful IV KRAS

Figure S4. Effect of different co-mutations and *TP53* mutation variants on malignant program activity and signaling entropy.

A. Comparison of malignant program scores and *Hif1a* expression between K (blue) *vs.* KP (red) LUAD mouse model scRNA-seq data malignant cells. *P*-values were calculated using a Mann-Whitney-Wilcoxon test.

B. Distribution of entropy scores across adjacent normal (grey), *TP53*^{WT} (blue), and different *TP53*^{mut} categories (reds) in bulk RNA-seq LUAD tumor samples from TCGA. *P*-values were calculated using a Mann-Whitney-Wilcoxon test.

C. Distribution of entropy scores across *TP53*^{WT} (blue), and different *TP53*^{mut} single and co-mutations (reds) in bulk RNA-seq LUAD tumor samples from TCGA. *P*-values were calculated using a Mann-Whitney-Wilcoxon test.

D. Heatmap of mean entropy score for malignant subsets in scRNA-seq data.

E. Stacked bar plot of proportion of HPCS cells from K and KP mice, classified into different malignant subsets from this study.

F. Comparison of entropy scores between *KRAS*^{WT} and different *KRAS*^{mut} impact categories in A549 cells.

Figure S5

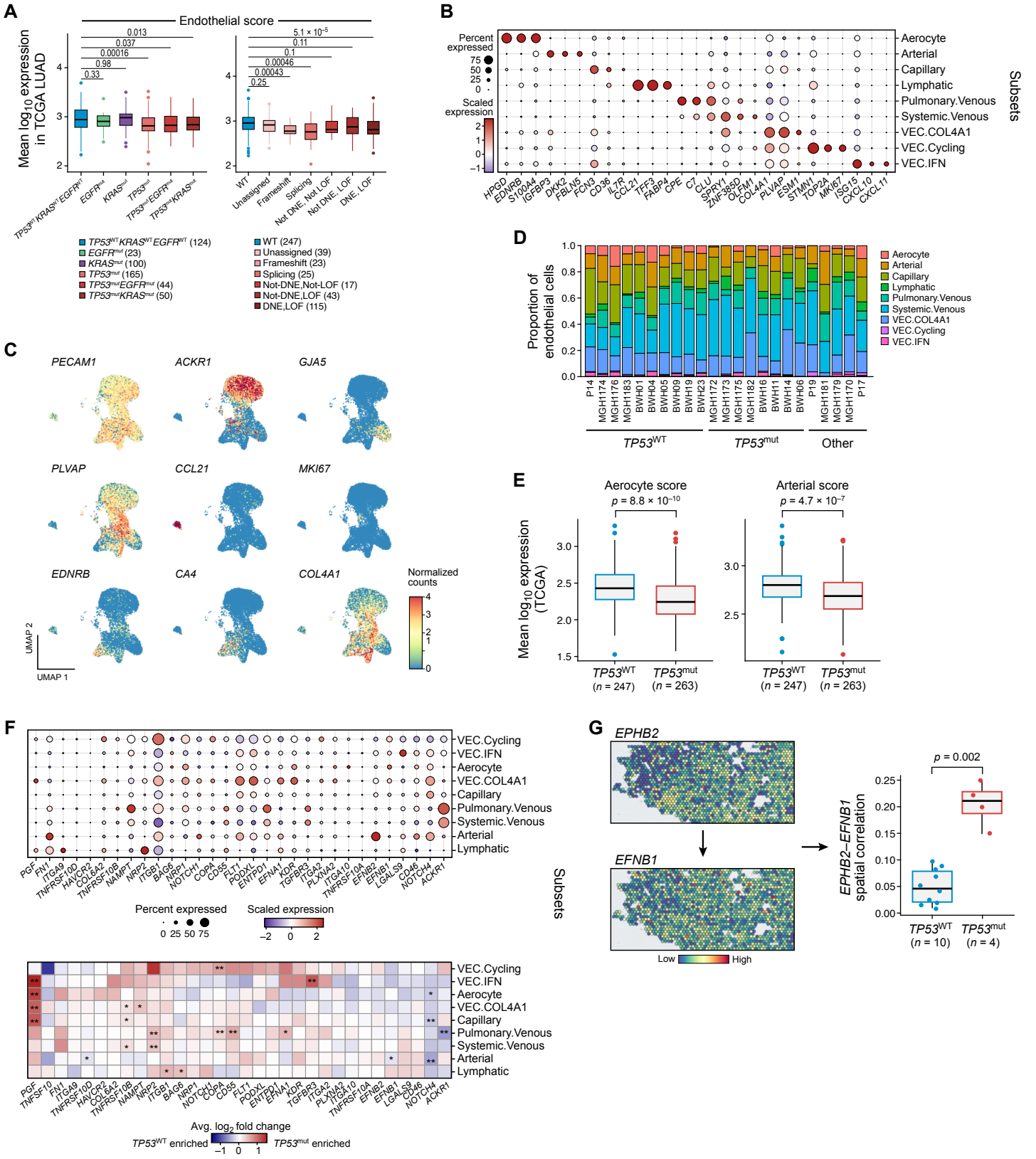
A
Endothelial score
Mean log10 expression in TCGA LUAD
0.013
0.037
0.00016
0.98
0.33
5.1 × 10−5
0.11
0.1
0.00046
0.00043
0.25
TP53WT KRASWT EGFRWT
EGFRmut
KRASmut
TP53mut
TP53mut EGFRmut
TP53mut KRASmut
WT
Unassigned
Frameshift
Splicing
Not DNE, Not LOF
Not DNE, LOF
DNE, LOF
TP53WT KRASWT EGFRWT (124)
EGFRmut (23)
KRASmut (100)
TP53mut (165)
TP53mut EGFRmut (44)
TP53mut KRASmut (50)
WT (247)
Unassigned (39)
Frameshift (23)
Splicing (25)
Not-DNE,Not-LOF (17)
Not-DNE,LOF (43)
DNE,LOF (115)
B
Percent expressed
Scaled expression
Aerocyte
Arterial
Capillary
Lymphatic
Pulmonary.Venous
Systemic.Venous
VEC.COL4A1
VEC.Cycling
VEC.IFN
Subsets
HPGD
EDNRB
S100A4
IGFBP3
DKK2
FBLN5
FCN3
CD36
IL7R
CCL21
TFF3
FABP4
CPE
C7
CLU
SPRY1
ZNF385D
OLFM1
COL4A1
PLVAP
ESM1
STMN1
TOP2A
MKI67
ISG15
CXCL10
CXCL11
C
PECAM1
ACKR1
GJA5
PLVAP
CCL21
MKI67
EDNRB
CA4
COL4A1
Normalized counts
UMAP 2
UMAP 1
D
Proportion of endothelial cells
Aerocyte
Arterial
Capillary
Lymphatic
Pulmonary.Venous
Systemic.Venous
VEC.COL4A1
VEC.Cycling
VEC.IFN
P14
MGH174
MGH176
MGH183
BWH01
BWH04
BWH05
BWH09
BWH19
BWH23
MGH172
MGH173
MGH175
MGH182
BWH16
BWH11
BWH14
BWH06
P19
MGH181
MGH179
MGH170
P17
TP53WT
TP53mut
Other
E
Aerocyte score
p = 8.8 × 10−10
Arterial score
p = 4.7 × 10−7
Mean log10 expression (TCGA)
TP53WT (n = 247)
TP53mut (n = 263)
F
VEC.Cycling
VEC.IFN
Aerocyte
VEC.COL4A1
Capillary
Pulmonary.Venous
Systemic.Venous
Arterial
Lymphatic
Percent expressed
Scaled expression
Avg. log2 fold change
TP53WT enriched
TP53mut enriched
G
EPHB2
EFNB1
Subsets
Low
High
p = 0.002
EPHB2–EFNB1 spatial correlation
TP53WT (n = 10)
TP53mut (n = 4)

Figure S5. Characterizing endothelial subsets and their differential roles in *TP53*^{WT} and *TP53*^{mut} LUAD.

A. Left: Comparison of mean $\log_{10}$ endothelial marker expression between *TP53*^{WT} and different single- or co-mutations of *EGFR*, *KRAS*, and *TP53* in bulk RNA-seq data from TCGA tumor samples. Right: Comparison of mean $\log_{10}$ endothelial marker expression between *TP53*^{WT} and different dominant-negative/loss of function *TP53*^{mut} variant categories in bulk RNA-seq LUAD tumor samples from TCGA. Mann-Whitney-Wilcoxon test was used for computing *p*-values.

B. Dot plot showing the expression of three marker genes for each annotated endothelial subset.

C. Feature plots of the expression of nine selected endothelial subset markers.

D. Stacked bar plot displaying the cell composition of endothelial subsets for each scRNA-seq patient sample.

E. Average $\log_{10}$ expression of aerocytes (left) and arterial (right) markers derived from the scRNA-seq data for *TP53*^{WT} (blue) and *TP53*^{mut} (red) LUAD bulk RNA-seq tumor samples from TCGA. *P*-values were calculated using a two-tailed multiple regression t-test.

F. Top: Dot plot of genes involved in ligand-receptor interactions displayed in **Figure 3C,** Bottom: Pseudobulk differential expression analysis performed on the same genes comparing *TP53*^{mut} *vs*. *TP53*^{WT} tumor samples. Mann-Whitney-Wilcoxon test *p*-values were calculated. Red indicates enrichment in *TP53*^{mut} and blue enrichment in *TP53*^{WT} tumor samples.

G. Left: Spatial expression of *EFHB2* and *EFNB1* in a representative *TP53*^{mut} ST sample. Right: Box plot comparing spatial correlation of *EFHB2* and *EFNB1* between *TP53*^{mut} and *TP53*^{WT} ST tumor samples. Mann-Whitney-Wilcoxon test *p*-values were computed.

Figure S6

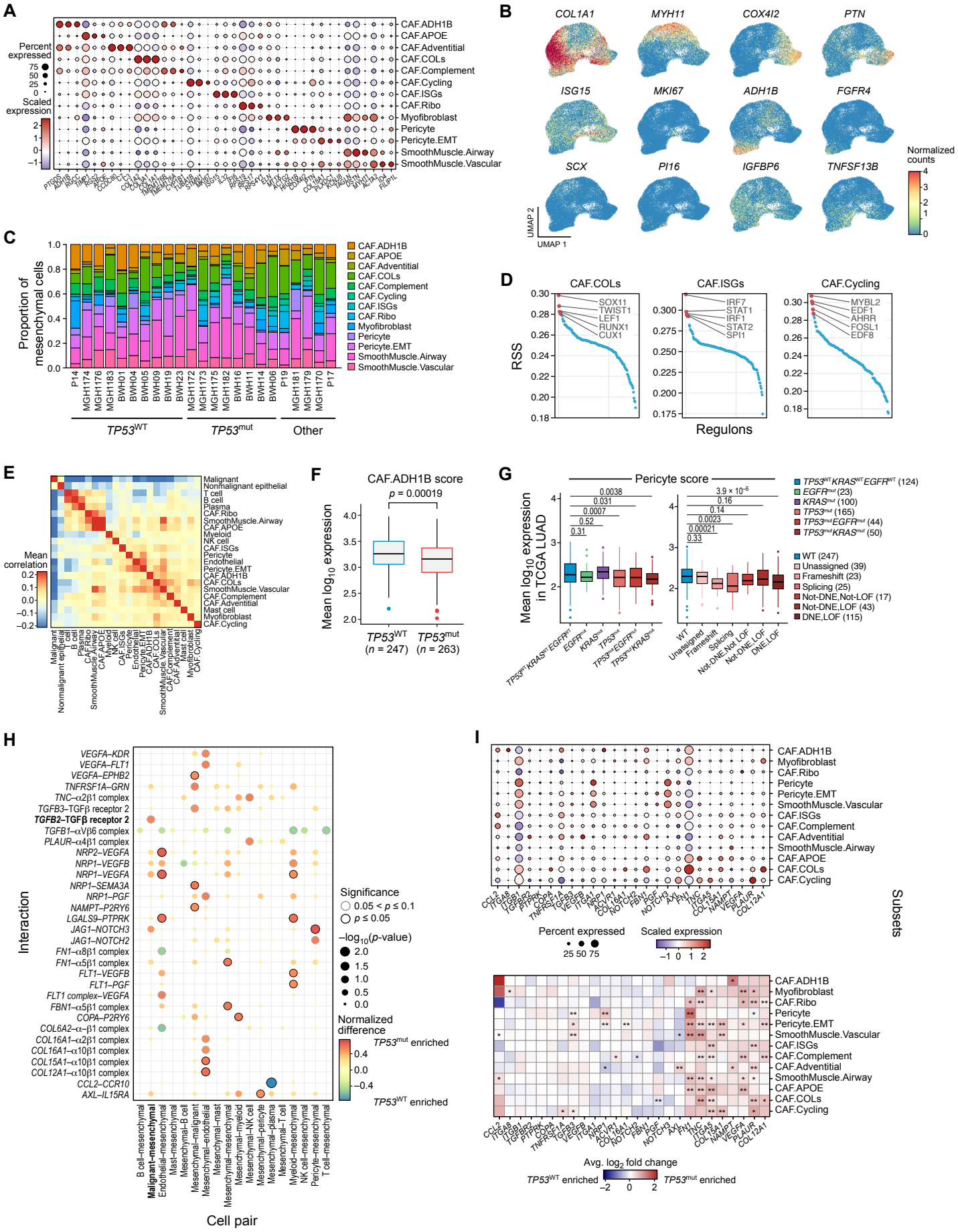
A
Percent expressed
Scaled expression
CAF.ADH1B
CAF.APOE
CAF.Adventitial
CAF.COLs
CAF.Complement
CAF.Cycling
CAF.ISGs
CAF.Ribo
Myofibroblast
Pericyte
Pericyte.EMT
SmoothMuscle.Airway
SmoothMuscle.Vascular
B
COL1A1
MYH11
COX4I2
PTN
ISG15
MKI67
ADH1B
FGFR4
SCX
PI16
IGFBP6
TNFSF13B
Normalized counts
UMAP 2
UMAP 1
C
Proportion of mesenchymal cells
TP53WT
TP53mut
Other
D
CAF.COLs
CAF.ISGs
CAF.Cycling
RSS
Regulons
E
Mean correlation
F
CAF.ADH1B score
p = 0.00019
Mean log10 expression
TP53WT (n = 247)
TP53mut (n = 263)
G
Pericyte score
Mean log10 expression in TCGA LUAD
H
Interaction
Cell pair
Significance
Normalized difference
I
Subsets
Percent expressed
Scaled expression
Avg. log2 fold change

Figure S6. Mesenchymal cell compartment differences between *TP53*^{mut} and *TP53*^{WT} LUAD.

A. Dot plot showing the expression of three marker genes for each annotated mesenchymal subset.

B. Feature plot showing the expression of 12 selected mesenchymal subset markers.

C. Stacked bar plot displaying the cell composition of mesenchymal subsets for each patient sample.

D. Regulon specificity scores (RSS) plotted for three different mesenchymal subsets of interest.

E. Spatial correlation (colocalization) of mesenchymal subsets with other broad cell classes across all ST tumor samples. Red indicates high colocalization and blue low colocalization between pairs of cell subsets.

F. Average $\log_{10}$ expression of CAF.ADH1B markers derived from the scRNA-seq data for *TP53*^{WT} (blue) and *TP53*^{mut} (red) LUAD bulk RNA-seq tumor samples from TCGA. *P*-values were calculated using a two-tailed multiple regression t-test.

G. Left: Comparison of mean $\log_{10}$ pericyte marker expression between *TP53*^{WT} and different single- or co-mutations of *EGFR*, *KRAS*, and *TP53* in bulk RNA-seq data from TCGA tumor samples. Right: Comparison of mean $\log_{10}$ pericyte marker expression between *TP53*^{WT} and different dominant-negative/loss of function *TP53*^{mut} variant categories in bulk RNA-seq LUAD tumor samples from TCGA. Mann-Whitney-Wilcoxon test was used for computing *p*-values.

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

I. Top: Dot plot of genes involved in ligand-receptor interactions displayed in **Figure S3H.** Bottom: Pseudobulk differential expression analysis performed on the same genes comparing

TP53^{mut} *vs*. *TP53*^{WT} tumor samples. Mann-Whitney-Wilcoxon test *p*-values were calculated. Red indicates enrichment in *TP53*^{mut} and blue enrichment in *TP53*^{WT} tumor samples.

Figure S7

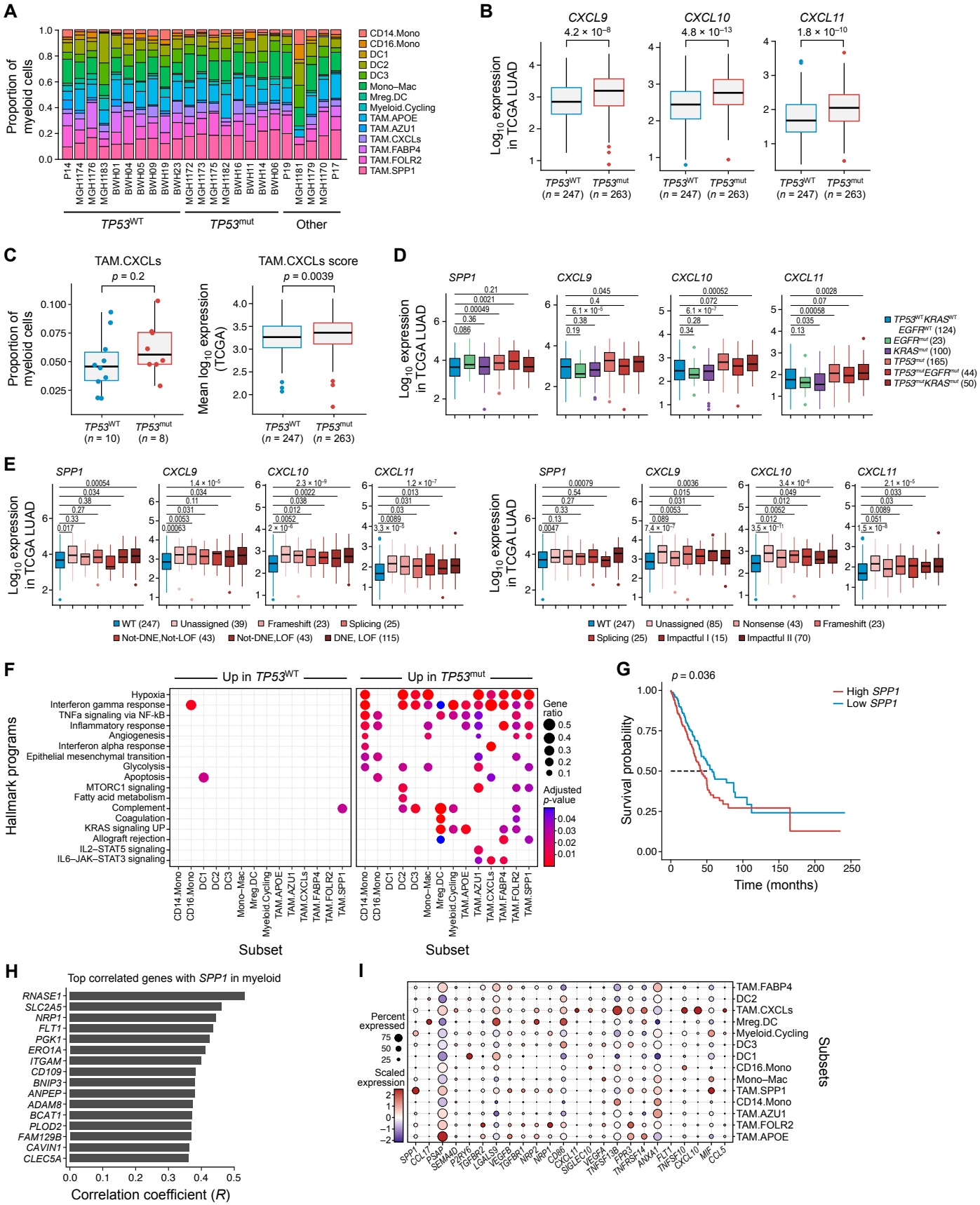
A
Proportion of myeloid cells
CD14.Mono
CD16.Mono
DC1
DC2
DC3
Mono–Mac
Mreg.DC
Myeloid.Cycling
TAM.APOE
TAM.AZU1
TAM.CXCLs
TAM.FABP4
TAM.FOLR2
TAM.SPP1
TP53WT
TP53mut
Other
B
CXCL9
4.2 × 10−8
CXCL10
4.8 × 10−13
CXCL11
1.8 × 10−10
Log10 expression in TCGA LUAD
TP53WT (n = 247)
TP53mut (n = 263)
C
TAM.CXCLs
p = 0.2
Proportion of myeloid cells
TP53WT (n = 10)
TP53mut (n = 8)
TAM.CXCLs score
p = 0.0039
Mean log10 expression (TCGA)
TP53WT (n = 247)
TP53mut (n = 263)
D
SPP1
CXCL9
CXCL10
CXCL11
Log10 expression in TCGA LUAD
TP53WT KRASWT EGFRWT (124)
EGFRmut (23)
KRASmut (100)
TP53mut (165)
TP53mut EGFRmut (44)
TP53mut KRASmut (50)
E
SPP1
CXCL9
CXCL10
CXCL11
Log10 expression in TCGA LUAD
WT (247)
Unassigned (39)
Frameshift (23)
Splicing (25)
Not-DNE,Not-LOF (43)
Not-DNE,LOF (43)
DNE, LOF (115)
WT (247)
Unassigned (85)
Nonsense (43)
Frameshift (23)
Splicing (25)
Impactful I (15)
Impactful II (70)
F
Up in TP53WT
Up in TP53mut
Hallmark programs
Hypoxia
Interferon gamma response
TNFa signaling via NF-kB
Inflammatory response
Angiogenesis
Interferon alpha response
Epithelial mesenchymal transition
Glycolysis
Apoptosis
MTORC1 signaling
Fatty acid metabolism
Complement
Coagulation
KRAS signaling UP
Allograft rejection
IL2–STAT5 signaling
IL6–JAK–STAT3 signaling
Gene ratio
Adjusted p-value
Subset
G
p = 0.036
High SPP1
Low SPP1
Survival probability
Time (months)
H
Top correlated genes with SPP1 in myeloid
RNASE1
SLC2A5
NRP1
FLT1
PGK1
ERO1A
ITGAM
CD109
BNIP3
ANPEP
ADAM8
BCAT1
PLOD2
FAM129B
CAVIN1
CLEC5A
Correlation coefficient (R)
I
Percent expressed
Scaled expression
Subsets
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
FPR3
TNFRSF14
ANXA1
FLT1
TNFSF10
CXCL10
MIF
CCL5

Figure S7. Characterization of myeloid subsets and expression of ligand-receptor genes.

A. Stacked bar plot displaying the cell composition of myeloid subsets for each patient sample.

B. Comparison of *CXCL9/10/11* expression in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD tumor samples using bulk RNA-seq data from TCGA.

C. Left: Proportion of TAM.CXCLs subset relative to all myeloid cells in *TP53*^{WT} (blue) *TP53*^{mut} (red) LUAD scRNA-seq tumor samples. Right: Average $\log_{10}$ expression of TAM.CXCLs markers in LUAD bulk RNA-seq tumor samples from TCGA. *P*-values were calculated using a Mann-Whitney-Wilcoxon test (scRNA-seq, left) or a two-tailed multiple regression t-test (TCGA, right).

D. Comparison of *SPP1*, *CXCL9*/*10*/*11* gene expression among *TP53*^{WT} and different single- or co-mutations in *EGFR*, *KRAS*, and *TP53.*

E. Comparison of gene expression among dominant-negative/loss of function (left), and impact categories (right) of *TP53*^{mut} LUAD tumor samples in TCGA. *P*-values were calculated using a Mann-Whitney-Wilcoxon test.

F. Dot plot of gene set enrichments using the Hallmark database for the 40 most differentially expressed genes in each myeloid subsets, comparing genes up in *TP53*^{WT} (left) and up in *TP53*^{mut} (right).

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

H. Horizontal bar plot showing the most correlated genes with *SPP1* expression in myeloid cells, ranked by Pearson's correlation coefficient.

I. Dot plot showing expression of genes involved in ligand-receptor interactions shown in **Figure 4F.**

Figure S8

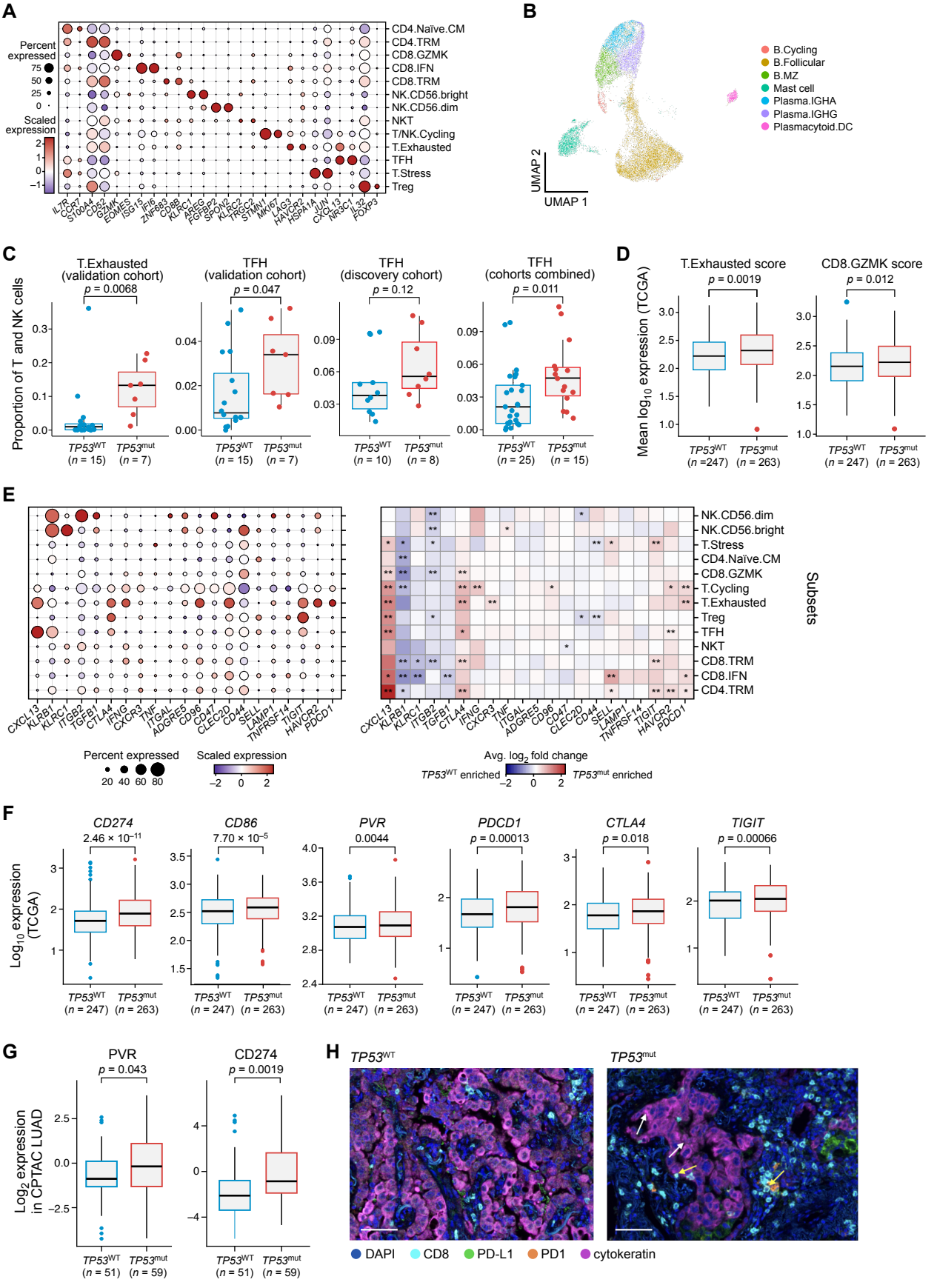
A
Percent expressed
Scaled expression
CD4.Naïve.CM
CD4.TRM
CD8.GZMK
CD8.IFN
CD8.TRM
NK.CD56.bright
NK.CD56.dim
NKT
T/NK.Cycling
T.Exhausted
TFH
T.Stress
Treg
B
B.Cycling
B.Follicular
B.MZ
Mast cell
Plasma.IGHA
Plasma.IGHG
Plasmacytoid.DC
UMAP 2
UMAP 1
C
T.Exhausted (validation cohort)
TFH (validation cohort)
TFH (discovery cohort)
TFH (cohorts combined)
Proportion of T and NK cells
D
T.Exhausted score
CD8.GZMK score
Mean log10 expression (TCGA)
E
Subsets
Percent expressed
Scaled expression
Avg. log2 fold change
TP53WT enriched
TP53mut enriched
F
CD274
CD86
PVR
PDCD1
CTLA4
TIGIT
Log10 expression (TCGA)
G
PVR
CD274
Log2 expression in CPTAC LUAD
H
TP53WT
TP53mut
DAPI
CD8
PD-L1
PD1
cytokeratin

Figure S8. Lymphoid immune checkpoint gene expression and spatial distribution.

A. Dot plot showing the expression of two representative marker genes for each annotated T and NK subset.

B. UMAP of 18,520 B, plasma, mast, and plasmacytoid DC cells, integrated across tumors and colored by annotated subset.

C. Comparison of T.Exhausted (left; validation cohort only) and TFH cell proportions (right; discovery, validation, and combined cohorts) relative to all T and NK cells between *TP53*^{mut} and *TP53*^{WT} tumors profiled by scRNA-seq. *P*-values were calculated using a Mann-Whitney-Wilcoxon test.

D. Average $\log_{10}$ expression of specific T.Exhausted, and CD8.GZMK markers derived from the scRNA-seq data in *TP53*^{WT} (blue) and *TP53*^{mut} (red) LUAD bulk RNA-seq tumor samples from TCGA. *P*-values were calculated using a two-tailed multiple regression t-test.

E. Left: Dot plot of genes involved in ligand-receptor analysis displayed in **Figure 5D**. Right: Pseudobulk differential expression analysis performed on the same genes comparing *TP53*^{mut} *vs*. *TP53*^{WT} tumor samples. Mann-Whitney-Wilcoxon test *p*-values were calculated. Red indicates enrichment in *TP53*^{mut} and blue enrichment in *TP53*^{WT} tumor samples.

F. Log_{10} expression of *CD274*, *CD86*, *PVR*, *PDCD1, CTLA4,* and *TIGIT* in *TP53*^{WT} (blue) and *TP53*^{mut} (red) LUAD bulk RNA-seq tumor samples from TCGA. *P*-values were calculated using a two-tailed multiple regression t-test.

G. Boxplots of PVR and CD274 (PD-L1) protein expression in CPTAC tumor samples split by *TP53* mutational status. Mann-Whitney-Wilcoxon test *p*-values were computed.

H. Representative mIF images of additional *TP53*^{WT} (BWH04) and *TP53*^{mut} (BWH11) tumor stains. Yellow arrows correspond to PD1^{+}CD8^{+} cells and white arrows correspond to PD-L1^{+}cytokeratin^{+} cells. Scale bar, 50 µm.

Figure S9

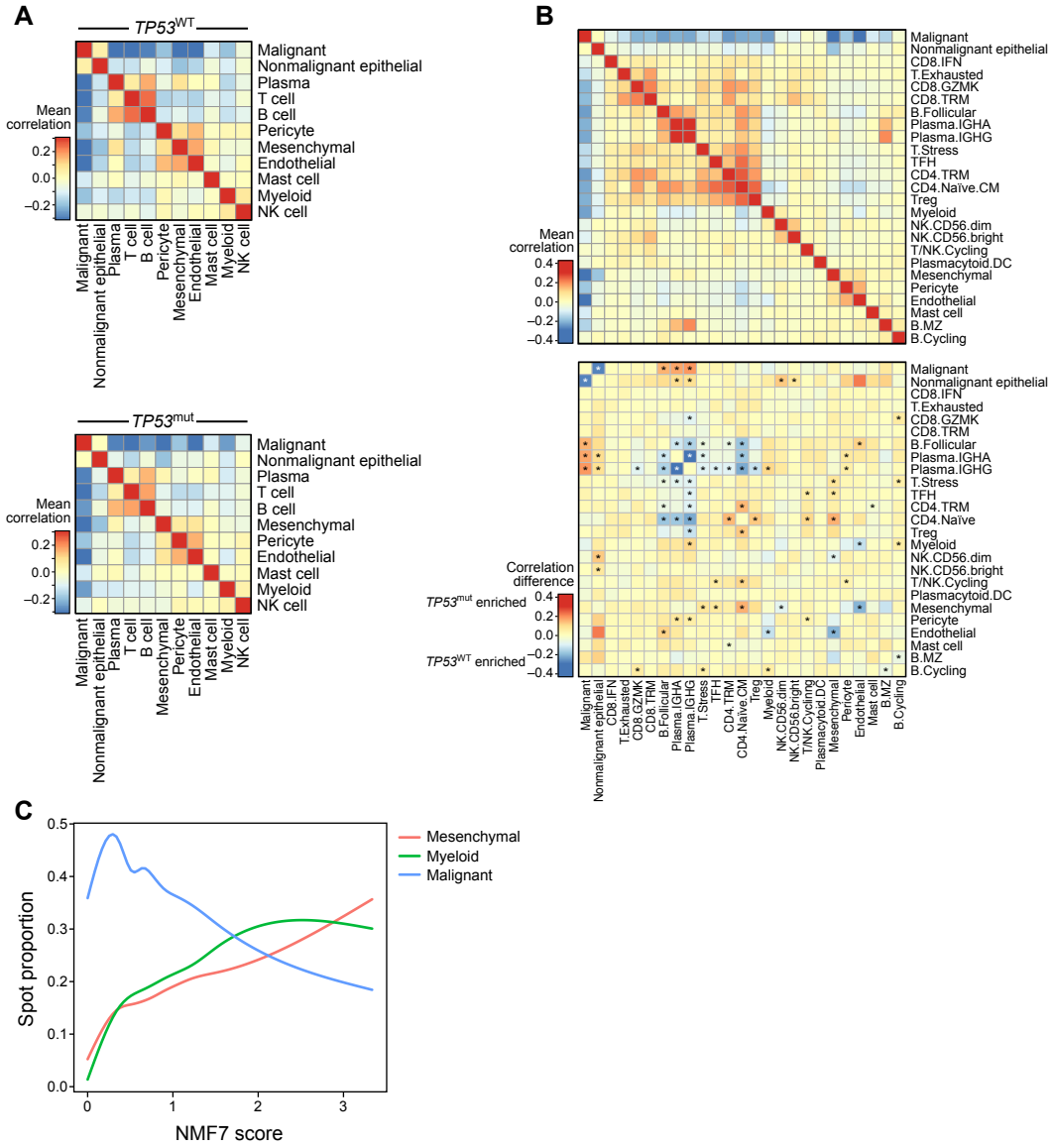
A
TP53WT
Mean correlation
0.2
0.0
−0.2
Malignant
Nonmalignant epithelial
Plasma
T cell
B cell
Pericyte
Mesenchymal
Endothelial
Mast cell
Myeloid
NK cell
TP53mut
B
Mean correlation
0.4
0.2
0.0
−0.2
−0.4
Malignant
Nonmalignant epithelial
CD8.IFN
T.Exhausted
CD8.GZMK
CD8.TRM
B.Follicular
Plasma.IGHA
Plasma.IGHG
T.Stress
TFH
CD4.TRM
CD4.Naïve.CM
CD4.Naïve
Treg
Myeloid
NK.CD56.dim
NK.CD56.bright
T/NK.Cycling
Plasmacytoid.DC
Mesenchymal
Pericyte
Endothelial
Mast cell
B.MZ
B.Cycling
Correlation difference
TP53mut enriched
TP53WT enriched
C
Spot proportion
NMF7 score
Mesenchymal
Myeloid
Malignant

Figure S9. Spatial organization and cellular landmarks in *TP53*^{mut} and *TP53*^{WT} LUAD.

A. Mean spatial correlation of broad cell classes across ST spots averaged across *TP53*^{WT} (top) and *TP53*^{mut} (bottom) samples.

B. Mean (top) and differential (bottom) spatial correlation heatmap of lymphoid (B, plasma, T, NK) subsets with other broad cell classes between *TP53*^{mut} and *TP53*^{WT}. Red indicates either high overall colocalization (top) or colocalization enrichment in *TP53*^{mut} (bottom) while blue represents either low overall colocalization (top) or colocalization enrichment in *TP53*^{WT} (bottom) tumor samples between pairs of cell subsets.

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

Table S1. Clinical data for patient cohort

Table S2. High level cell class markers and subset markers

Table S3. Clinical data for validation cohort