AbdulJabbar, K. et al. (2020) Geospatial immune variability illuminates differential evolution of lung adenocarcinoma. Nature Medicine, 26(7), pp. 1054-1062. (doi: 10.1038/s41591-020-0900-x)

1 **Geospatial immune variability illuminates differential evolution of**

2 **lung adenocarcinoma**

3

4 Khalid AbdulJabbar[1,2^], Shan E Ahmed Raza[1,2^], Rachel Rosenthal[3,4#], Mariam Jamal-Hanjani[3,5#],

5 Selvaraju Veeriah[3,4#], Ayse Akarca[6], Tom Lund[7], David A. Moore[3,6], Roberto Salgado[8,9], Maise Al

6 Bakir[4], Luis Zapata[1,2], Crispin T. Hiley[3,4], Leah Officer[10], Marco Sereno[11], Claire Rachel Smith[11],

7 Sherene Loi[9], Allan Hackshaw[12], Teresa Marafioti[6], Sergio A. Quezada[13], Nicholas McGranahan[3,14],

8 John Le Quesne[10,11,15*], TRACERx consortium, Charles Swanton[3,4,5*] and Yinyin Yuan[1,2*]

9

10 ^: Joint first authors

11 #: Joint second authors

12 *: Joint corresponding authors

13

14 1Centre for Evolution and Cancer, The Institute of Cancer Research, London, UK

15 2Division of Molecular Pathology, The Institute of Cancer Research, London, UK

16 3Cancer Research UK Lung Cancer Centre of Excellence, University College London Cancer Institute,

17 London, UK

18 4Cancer Evolution and Genome Instability Laboratory, The Francis Crick Institute, London, UK

19 5Department of Medical Oncology, University College London Hospitals NHS Foundation Trust, 20 London, UK

21 6Department of Cellular Pathology, University College London, University College Hospital, London,

22 UK

23 7Translational Immune Oncology Group, Centre for Molecular Medicine, Royal Marsden Hospital

24 NHS Trust, London, UK

25 8Department of Pathology, GZA-ZNA-Ziekenhuizen, Antwerp, Belgium

26 9Division of Research, Peter MacCallum Cancer Centre, University of Melbourne, Melbourne, 27 Victoria, Australia

28 10MRC Toxicology Unit, Lancaster Road, University of Cambridge, Leicester, UK

29 11Leicester Cancer Research Centre, University of Leicester, Leicester, UK

30 12Cancer Research UK & University College London Cancer Trials Centre, University College London,

31 London, UK

32 13Cancer Immunology Unit, University College London Cancer Institute, London, UK

33 14Cancer Genome Evolution Research Group, University College London Cancer Institute, University

34 College London, London, UK

35 15Glenfield Hospital, University Hospitals Leicester NHS Trust, Groby Road, Leicester, UK

Abstract

Remarkable progress in molecular analyses has improved our understanding of the evolution of cancer cells towards immune escape[1–5]. However, the spatial configurations of immune and stromal cells, which may shed light on the evolution of immune escape across tumor geographical locations, remain unaddressed. We integrated multi-region exome and RNA-seq data with spatial histology mapped by deep learning in 100 non-small cell lung cancer (NSCLC) patients from the TRAcking Cancer Evolution through Therapy (Rx) (TRACERx) cohort[6]. Cancer subclones derived from immune cold regions were more closely related in mutation space, diversifying more recently than subclones from immune hot regions. In TRACERx and in an independent multi-sample cohort of 970 lung adenocarcinoma (LUAD) patients, the number of immune cold regions significantly correlated with risk of relapse, independently of tumor size, stage and number of samples per patient. In LUAD, but not lung squamous cell carcinoma (LUSC), geometrical irregularity and complexity of the cancer-stromal cell interface significantly increased in tumor regions without disruption of antigen presentation. Decreased lymphocyte accumulation in adjacent stroma was observed in tumors with low clonal neoantigen burden. Collectively, immune geospatial variability elucidates tumor ecological constraints that may shape the emergence of immune evading subclones and aggressive clinical phenotypes.

56 **Main Text**

57

58 Using an artificial intelligence framework, we developed a generalizable deep learning

59 pipeline to spatially profile immune infiltration and discover tumor topological

60 determinants of immunosuppression in digital pathology. Convolutional neural networks

61 were tailored for the analysis of NSCLC morphology using diverse histology samples in the

62 multi-region TRACERx 100 cohort[6] to avoid overfitting (Methods). This approach enabled

63 the spatial mapping of cancer cells, lymphocytes, stromal cells (fibroblasts and endothelial

64 cells), and an "other" cell class (macrophages, pneumocytes and non-identifiable cells) in

65 hematoxylin & eosin (H&E)-stained images (275 tumor regions from 85 patients and 100

66 diagnostic slides from all patients, Fig. 1a-c, CONSORT diagram Extended Data Fig. 1a-b,

67 Supplementary Table 1). T cell subsets were also identified in CD4/CD8/FOXP3

68 immunohistochemistry (IHC) images for all 100 diagnostic samples (Fig. 1d).

69

70 This pipeline for H&E analysis exhibited high accuracy and consistency compared with five

71 orthogonal data types within TRACERx, including DNA-seq, RNA-seq, IHC, 5,951 single-cell

72 annotations by pathologists (balanced accuracy, as an average of specificity and sensitivity

73 = 0.932), and pathology tumor-infiltrating lymphocyte (TIL) estimates following the

74 guidelines developed by the International Immuno-Oncology Biomarker Working Group[7]

75 (Extended Data Fig. 2, Supplementary Table 2). The Leicester Archival Thoracic Tumor

76 Investigatory Cohort[8] (LATTICe-A, Extended Data Fig. 1c-d), a retrospective study of 970

77 resected LUAD patients that included H&E sections from all diagnostic tumor blocks with a

78 median of four samples per tumor, was used for independent validation. The pipeline's

79 generalizability was supported using 5,082 pathologists' single-cell annotations (balanced

80 accuracy = 0.913), and virtual integration of IHC and H&E images generated from the same

81 slides (Fig. 1e-h, Extended Data Fig. 2e-g, Supplementary Table 3). Using this unbiased

82 scalable approach, immune infiltration was quantified as the percentage of all cells that

83 were lymphocytes in each H&E image.

84

85 High geospatial immune variability between tumor regions within the same patients was

86 revealed (Fig. 2a-b), which did not reflect associations with pathological stage (Extended

87 Data Fig. 3). To differentiate highly from poorly immune infiltrated tumor regions, regions

88 containing a lymphocyte percentage greater than a quarter standard deviation above the

89 median lymphocyte percentage were classified as immune hot, and regions containing a

90 lymphocyte percentage below a quarter standard deviation of the median were classified

91 as immune cold. The remaining 20% were classified as intermediate (Fig. 2b). Subsequent

92 results were tested on four more classification schemes based on the standard deviation

93 to ensure that results derived from this classification were not contingent upon choice of

94 thresholds used (Extended Data Fig. 4). Significant difference in pathology TIL estimates

95 was observed between immune hot and cold regions (P = 4.6x10^{-8}, Extended Data
96 Fig. 5a). Significantly higher levels of RNA-seq estimated immune infiltrate[1], particularly for
97 immune activation subsets, were consistently observed in immune hot compared to cold
98 regions, supporting the validity of histology-based immune classification (Fig. 2c-d). We
99 next directly compared our immune hot and cold regional classification (excluding
100 intermediate regions) against RNA-seq-based[1] classifications (= 109 regions with histology
101 and RNA-seq data). 78 out of 109 regions were in agreement (Fisher's exact test for
102 overlap: P = 7.8x10^{-6}, Extended Data Fig. 5b). Regions with discrepant classification (
103 = 31) had significantly higher spatial heterogeneity of lymphocyte distribution compared
104 to regions concordant between the two methods (P = 0.01, Extended Data Fig. 5c),
105 suggesting spatial intratumor heterogeneity could contribute towards the discrepancy,
106 since the different data types were derived from adjacent sections of the same tumor
107 blocks.

108

109 Ecological selection pressures drive genetic divergence[9,10]. To determine if cancer genetic
110 divergence differs according to immune context, we calculated the genomic distance as
111 the Euclidean distance of subclonal mutations for each pair of tumor regions with the
112 same immune phenotype in a patient. We observed significantly lower genomic distance,
113 indicating more shared subclonal mutations, for pairs of immune cold regions than for
114 pairs of immune hot regions in LUAD (Fig. 3a, Extended Data Fig. 4b, P < 0.005 for
115 all immune classification schemes), but not in LUSC (Extended Data Fig. 6a). In LUAD but
116 not LUSC, analysis of immune phenotypes mapped onto the phylogenetic trees[6] revealed
117 that dominant clones (cancer cell fraction ≥ 75%, see Methods) in pairs of cold regions
118 were more closely related on the phylogenetic tree, compared to dominant clones in pairs
119 of immune hot regions (Fig. 3b). Moreover, dominant clones in hot regions almost always
120 diversified at the most recent common ancestor of the tree (13/15, 87%, Fig. 3c), in
121 contrast no such preference was observed in immune cold regions (11/23, 48%).

122

123 We investigated the impact of immune context on disease-free survival. Tumors with high
124 number of immune cold regions were at significantly increased risk of relapse that was
125 independent of the total number of regions sampled, tumor size and stage in both
126 histology types in TRACERx (Fig. 3d-e, Extended Data Fig. 6c-h). This association with
127 disease-free survival was also significant using the number of immune low regions as
128 estimated by RNAseq[1] in 64 TRACERx tumors with available RNA-seq data (P = 0.002,
129 Extended Data Fig. 6b). Following the genomic findings in LUAD, we sought to validate this
130 in 970 LUAD patients in the multi-sample LATTICe-A cohort, confirming the prognostic
131 value of immune cold sample count, that was also independent of the number of samples
132 per patient, tumor size and stage (Fig. 3f-g, Extended Data Fig. 6c-e). In both cohorts, the
133 number of immune cold samples per patient correlated with relapse, more significantly
134 than any other immune feature generated using deep learning, including the average and

variability of lymphocyte percentage per tumor, number of immune hot regions, proportion of immune cold regions to the number of regions sampled, as well as CD8[+] cell percentage or CD8[+] to CD4[+]FOXP3[+] ratio in TRACERx diagnostic slides (Extended Data Fig. 6e).

Studies have revealed immunosuppressive fibroblast subsets localizing to the boundary of tumor nests possibly contribute to T cell exclusion[11–13]. Therefore, we hypothesized that increased cancer-stroma physical contact may reflect stroma-modulated inhibition of antitumor immune responses[14–17]. To measure the physical contact between cancer and stromal cells (the majority being fibroblasts) identified by image analysis, we developed a spatial measure, using fractal dimension to quantify the geographical irregularity and complexity of the cancer-stromal cell interface (Methods, Fig. 4a, Extended Data Fig. 7a,b,e). Within the same tissue space, higher fractal dimension of cancer-stromal cell interface suggests increased geometric irregularity and more extensive physical contact between tumor and stromal cells than samples with a smooth interface. For both histology types, fractal dimension was significantly higher in immune cold regions compared to immune hot regions (Fig. 4b, Extended Data Fig. 7c). Moreover, the difference in fractal dimension between immune cold and hot regions was more significant compared to the difference in stromal cell percentage (both histology types combined: R = 0.00036, effect size 0.49 for fractal dimension versus R = 0.018, effect size 0.38 for stromal cell percentage, Extended Data Fig. 7d), suggesting the importance of stromal cell geographical location rather than their quantity. This supports the hypothesis that the stroma-based inhibition of immune infiltration[17] may result from a specific topological pattern in the form of cancer-stroma engagement.

To understand the associations of stromal-mediated immunosuppression in the context of the genetic mechanisms of immune evasion, we related fractal dimension to dysfunction in antigen presentation through loss of heterozygosity at the human leukocyte antigen locus (HLA LOH), which has been identified as a potent immune escape mechanism[1,18]. A significantly higher fractal dimension was found in LUAD tumor regions with intact HLA alleles compared with regions harboring HLA LOH (Fig. 4c, Extended Data Fig. 7f). This was observed at the tumor level (see Methods for definition), independent of clonal neoantigen burden (R = 0.04, multivariate regression, Extended Data Fig. 7h), but was not observed in LUSC (Extended Data Fig. 7g, i).

Although clonal neoantigens have been associated with a cytotoxic immune response[19], the spatial distribution of lymphocytes in relation to clonal neoantigens remained unclear. To provide sufficient spatial context for analysis of cell distribution, whole-section TRACERx diagnostic H&E images, typically 10x larger than the regional samples, were used. To test the relationship between lymphocyte spatial distribution and clonal neoantigens,

175 we leveraged an established method for lymphocyte spatial modeling[20]. Each lymphocyte
176 was classified into three distinct spatial compartments: intra-tumor, adjacent-to-tumor or
177 distaltumor, based on unsupervised modeling of cancer-lymphocyte proximity (Fig. 4d). In
178 LUAD, but not LUSC, clonal neoantigens[19] were found to be associated with a specific
179 immune spatial score to approximate pathology TIL estimates[7], defined as the ratio of
180 adjacenttumor lymphocytes to stromal cells in the diagnostic H&E samples (R =
181 0.0074, high clonal neoantigen defined as above median in LUAD, Fig. 4e; correlation as
182 continuous variables Rho = 0.37, R = 0.035 after multiple testing correction, Extended
183 Data Fig. 8a). By contrast, subclonal neoantigen burden did not correlate with any immune
184 score (Extended Data Fig. 8a), supporting the notion that clonal but not subclonal
185 neoantigens is associated with infiltration of cytotoxic T cells[19] adjacent to tumor nests.

186

187 To determine if there was an enrichment of a specific lymphocyte subpopulation within
188 the adjacent-tumor compartment in LUAD, we spatially aligned IHC to H&E in 10 samples
189 with the highest adjacent-tumor lymphocytes to stromal cell ratio, and projected IHC-
190 derived T cell subsets onto H&E images, thereby creating virtual staining of cells in the
191 H&E sections (Methods, Fig. 4f, Extended Data Fig. 8b-c). $CD4^+FOXP3^-$, $CD8^+$, and
192 $CD4^+FOXP3^+$ cells classified in IHC were projected onto a density map of cancer cell
193 distribution inferred from H&E, and were classified into adjacent-tumor, intra-tumor, and
194 distal-tumor compartments. In this limited dataset, a significant increase of the effector-
195 regulator balance defined by $CD8^+/CD4^+FOXP3^+$ cell ratio was observed in adjacent-tumor
196 stroma compared to the distal tumor compartment (Fig. 4g).

197

198 In summary, by training deep learning algorithms in diverse histology samples, we
199 demonstrated that digital pathology can provide accurate tools for defining the ecological
200 spatial context that may improve our understanding of cancer evolution and the immune
201 response. In TRACERx and LATTICe-A cohorts, LUAD tumors with increased immune cold
202 regions were at a significantly higher risk of cancer relapse, independent of total regions
203 sampled and immune phenotypes of other regions. Thus, even within a tumor that has on
204 average increased immune infiltration, if it contains regions classified as immune cold,
205 prognosis appears to be associated with the number of cold regions. Analysis of cancer
206 branched evolution within the ecological context of immune hot and cold regions revealed
207 a difference in the evolution history of cancer subclones in these regions, possibly as a
208 result of immunoediting. Based on this finding, we speculate that by identifying the
209 subclone where immunoediting is likely to have occurred, new drivers of immune evasion
210 may be elucidated.

211

212 Spatial histology data can extend our knowledge of the tumor microenvironment
213 topological configuration in relation to genetic alterations relevant to immune
214 surveillance, including HLA LOH and clonal neoantigens in LUAD (Extended Data Fig. 9).
215 Increased cancerstromal engagement as measured by fractal dimension may signal
216 physical constraints against T cell ingress. This is supported by previous studies in lung
217 cancer showing restriction of CD8[+] and CD4[+] T cell motility in dense stromal extracellular
218 matrix areas around tumor epithelial cell regions which prevent them from entering tumor
219 islets[13]. Additionally, the association between specific spatial localization of lymphocytes in
220 tumoradjacent stroma and clonal neoantigens further support exploration of the role of
221 stromal cells in limiting tumor infiltration by T cells[14–17].

222

223 It will be imperative to validate our findings on a larger multi-region cohort of untreated
224 NSCLC tumors. Differences in our findings pertaining to LUAD and LUSC may reflect
225 differences in biology[21–23] and immune evasion mechanisms, including increased
226 prevalence of antigen presentation dysfunction (HLA transcriptional repression and HLA
227 LOH[1]) in LUSC. Other limitations include the lack of detailed staining using multiplexing
228 technologies[24–26] that could provide further insights into immune composition. However,
229 with advanced deep learning developments and detailed tumor phylogenetic data,
230 histology can be used to highlight fundamental immune contexture such as immune
231 exclusion and its topological determinants. These data illuminate the clinical significance
232 of immune cold regions that may reflect immune evading subclones, warranting further
233 investigation into mechanisms that could contribute to the spatial variability of immune
234 cells.

235

Figures legends

Figure 1. The computational pathology deep learning pipeline for dissecting heterogeneous NSCLC tumor microenvironment. a. Histology sample generation in Lung TRACERx. To preserve morphology and generate good quality histology, samples from the same tumor regional frozen blocks specifically collected for TRACERx and generated molecular data[1,6] were re-embedded in formalin fixed paraffin (FFPE). From these, H&Estained tumor section slides were generated. In addition, H&E section and triplex CD4/CD8/FOXP3 IHC slides were also generated from diagnostic blocks that represent clinical standard sampling. b. Our multistage deep learning pipeline consists of three key stages: fully automated tissue segmentation, single-cell detection and classification. The final output is shown as an image with all cells identified. For more details, please see the 'Training the deep learning pipeline' section of the Methods. c. Illustrative 3-dimensional distribution of input image patches in the feature space learned by the convolutional neural networks, using Principal Component Analysis. The feature clusters were pseudo-colored to display segregation for four cell types in H&E, and **d** CD8$^+$, CD4$^+$FOXP3$^+$, CD4$^+$FOXP3$^-$ and "other" cell class (hematoxylin cells) in IHC, respectively. **e**. The deep learning single-cell classification model was trained using expert pathology annotations from a variety of TRACERx samples (diagnostic, regional, TMA). The trained model was then applied to the remaining TRACERx samples (predominantly LUAD and LUSC) and the LATTICe-A cohort (only LUAD), identifying over 171 million cells in TRACERx and over 4.9 billion cells in LATTICe-A. WSI: whole-section image. **f**. Biological validation of the deep learning approach. H&E and IHC images generated from the same TMA slide were virtually integrated for comparison of H&E-based cell classification and cell type marker expression. For each marker, the experiment was conducted once using a single TMA (cores/patients = 48 TTF1; 38 CD45). Scale bars represent 100μm. **g-h**. Correlations between cancer/lymphocyte cell percentage determined by H&E and TTF1$^+$ (tumor marker)/CD45$^+$ (immune marker) cell percentage per LUAD image tiles of size 100μm^2 (= 100 TTF1; 83 CD45). The shading indicates 95% confidence interval.

Figure 2. Geospatial heterogeneity of lymphocytic infiltration in the TRACERx cohort. a. Representative examples of immune hot and immune cold multi-region H&E samples, scale bars represent 100μm. **b**. Each column represents a tumor, grouped by their histologic subtype (the "Other" group consists of adenosquamous carcinoma, large cell neuroendocrine carcinoma, pleomorphic carcinoma, and sarcomatoid carcinoma of pleomorphic type arising from adenocarcinoma). Tumor regions (illustrated as dots) were assigned to immune hot, immune cold, and intermediate phenotypes based on percentage of lymphocytes in all cells following H&E-based deep learning analysis. CD8$^+$/CD4$^+$FOXP3/CD4$^+$FOXP3$^+$ percentages based on automated analysis of the IHC

diagnostic samples are also shown. **c**. A heatmap showing gene expression patterns of 14 immune cell populations across tumor regions, each row represents a tumor region (= 142). The three clusters correspond to the proposed immune regional classification as shown in **b**. **d.** Significant enrichment of all immune cell populations in hot regions, as compared to cold regions, particularly for the immune activating cell subsets, including cytotoxic, B-cell, and natural killer cells (= 109 regions; 52 patients). A two-sided, non-parametric, unpaired, Wilcoxon signed-rank test was used for each box plot, all R-values were corrected for multiple comparisons. Thick horizontal lines indicate the median value; outliers are indicated by the extreme points; the first and third quantiles are represented by the box edges; and vertical lines indicate the error range.

Figure 3. Evolution of immune escape, and survival analysis in TRACERx and LATTICe-A. a. A box plot showing the difference in genomic distances for pairs of immune hot or immune cold regions within the same patients in LUAD (= 66 pairs). **b**. A box plot showing the difference in mutational distance between the dominant subclones in pairs of immune hot or immune cold regions via their last common ancestor in LUAD (= 23 immune cold pairs; 15 immune hot pairs). This distance was calculated by taking the furthest dominant clone (cancer cell fraction (CCF) ≥ 75%) from the trunk, and it remained significant when the dominant clone closest to the most recent common ancestor of each tree was considered (R = 0.02). **c**. Illustrative examples of tumor phylogenetic trees for a pair of immune hot and immune cold regions. Dominant subclones were labelled and their last common ancestor (annotated with arrows) was then identified. Minor (CCF < 75%) or undetected clones were neglected in this analysis. **d,e**. Kaplan-Meier curves illustrating the difference in disease-free survival according to the number of immune cold regions, dichotomized by the median value, in TRACERx (**d**) (LUAD and LUSC, = 79 patients, 249 regions) and LATTICe-A (**e**) (LUAD, = 970 patients, 4,324 samples). The same deep learning histology analysis and immune regional classification developed for TRACERx were applied directly to LATTICe-A. WSI: whole-section image. **f.** Forest plots showing multivariate Cox regression analyses in TRACERx (= 79 patients; LUAD and LUSC). Clonal neoantigens were dichotomized using the upper quartile, determined individually for LUAD and LUSC tumors[1]. **g.** Forest plots showing multivariate Cox regression analyses in LATTICe-A (= 651 LUAD patients with complete stage and smoking pack years data). For the patient subset with complete stage data but missing pack years information, the test remained significant (= 827, R < 0.001, HR = 1.4[1.1-1.9]). For statistical comparisons among groups, a two-sided, non-parametric, unpaired, Wilcoxon signed-rank test was used, unless stated otherwise.

Figure 4. Association of spatial histology with genetic alterations relevant to immune surveillance. a. An illustrative example of fractal dimension calculated by the box-counting

algorithm to quantify the geospatial complexity of the cancer cell-stromal cell interface. By examining boxes of decreasing sizes that contain both cancer and stromal cells, the box counting algorithm quantifies the rate at which the geometrical details of cancer-stromal interface develop at increasingly fine scales. Blue box illustrates the smallest box of 20µm by

20µm in size. Scale bar represent 100µm. An example of a fractal structure displaying geometrical self-similarity is shown below the panel. **b**. A box plot to illustrate the significant difference in fractal dimension between all TRACERx immune hot and cold regions (= 219).

c. A box plot showing a significant difference in fractal dimension between LUAD tumor regions (= 116) harboring an LOH event for class 1 HLA of any type versus regions that do not, adjusted for multiple comparisons with the remaining HLA type-specific tests (see Extended Data Fig. 7f). **d**. Illustration of the adjacent-tumor lymphocyte/stroma ratio inferred by spatial modeling of cancer cell density (contours) and lymphocyte classification into spatial compartments. Cell classification in IHC sample of the same block was shown for comparison. Scale bars represent 50µm. **e.** A box plot showing the difference in the adjacent-tumor lymphocyte/stroma ratio between high (≥ median) and low (< median) clonal neoantigens for all LUAD patients in TRACERx (= 61). **f**. Illustration of image registration to spatially align serial sections of H&E and IHC and generate a virtual composite map of T cell subset in the context of cancer/stroma density. T cell subsets classified in the IHC were projected onto the cancer density map inferred from H&E, so that they can be classified into adjacent-tumor, intra-tumor, and distal-tumor compartments. **g**. A box plot showing significantly higher ratio of $CD8^{+}$ to $CD4^{+}FOXP3^{+}$ cells in adjacent-tumor and intratumor lymphocytes compared with distal-tumor lymphocytes in registered LUAD image tiles (= 20 image tiles, using paired Wilcoxon test). For statistical comparisons among groups, a two-sided, non-parametric, unpaired, Wilcoxon signed-rank test was used, unless stated otherwise.

Main References

1.   Rosenthal, R. et al. Neoantigen-directed immune escape in lung cancer evolution. Nature 1 (2019). doi:10.1038/s41586-019-1032-7

2.   Morris, L. G. T. & Chan, T. A. Lung Cancer Evolution: What's Immunity Got to Do with It? Cancer Cell 35, 711–713 (2019).

3.   Morris, L. G. T. et al. Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. Oncotarget 7, 10051–10063 (2016).

352   4.   Milo, I. et al. The immune system profoundly restricts intratumor genetic
353   heterogeneity. Sci. Immunol. 3, (2018).

354   5.   Jia, Q. et al. Local mutational diversity drives intratumoral immune heterogeneity in
355   non-small cell lung cancer. Nat. Commun. 9, (2018).

356   6.   Jamal-Hanjani, M. et al. Tracking the Evolution of Non–Small-Cell Lung Cancer. N.
357   Engl. J. Med. 376, 2109–2121 (2017).

358   7.   Hendry, S. et al. Assessing Tumor-Infiltrating Lymphocytes in Solid Tumors. Adv.
359   Anat. Pathol. 24, 311–335 (2017).

360   8.   Moore, D. A. et al. In situ growth in early lung adenocarcinoma may represent
361   precursor growth or invasive clone outgrowth—a clinically relevant distinction.
362   Mod. Pathol.

363   1 (2019). doi:10.1038/s41379-019-0257-1

364   9.   Whittaker, K. A. & Rynearson, T. A. Evidence for environmental and ecological
365   selection in a microbe with no geographic limits to gene flow. Proc. Natl. Acad. Sci. U. S. A.
366   114, 2651–2656 (2017).

367   10.   Shafer, A. B. A. & Wolf, J. B. W. Widespread evidence for incipient ecological
368   speciation: A meta-analysis of isolation-by-ecology. Ecology Letters 16, 940–950 (2013).

369   11.   Costa, A. et al. Fibroblast Heterogeneity and Immunosuppressive Environment in
370   Human Breast Cancer. Cancer Cell 33, 463-479.e10 (2018).

371   12.   Öhlund, D. et al. Distinct populations of inflammatory fibroblasts and
372   myofibroblasts in pancreatic cancer. J. Exp. Med. 214, 579–596 (2017).

373   13.   Salmon, H. et al. Matrix architecture defines the preferential localization and
374   migration of T cells into the stroma of human lung tumors. J. Clin. Invest. 122, 899–910
375   (2012).

376   14.   Thomas, D. A. & Massagué, J. TGF-β directly targets cytotoxic T cell functions
377   during tumor evasion of immune surveillance. Cancer Cell 8, 369–380 (2005).

378   15.   Joyce, J. A. & Fearon, D. T. T cell exclusion, immune privilege, and the tumor
379   microenvironment. Science (80-.). 348, 74–80 (2015).

380   16.   Sorokin, L. The impact of the extracellular matrix on inflammation. Nat. Rev.
381   Immunol. 10, 712–723 (2010).

382   17.   Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer–immune set
383   point. Nature 541, 321–330 (2017).

384   18.   McGranahan, N. et al. Allele-Specific HLA Loss and Immune Escape in Lung Cancer
385   Evolution. Cell 171, 1259-1271.e11 (2017).

386   19.   McGranahan, N. et al. Clonal neoantigens elicit T cell immunoreactivity and
387   sensitivity to immune checkpoint blockade. Science (80-.). 351, (2016).

388   20.   Yuan, Y. Modelling the spatial heterogeneity and molecular correlates of
389   lymphocytic infiltration in triple-negative breast cancer. J. R. Soc. Interface 12, 20141153
390   (2015).

391   21.   Thomas, A., Liu, S. V., Subramaniam, D. S. & Giaccone, G. Refining the treatment of
392   NSCLC according to histological and molecular subtypes. Nat. Rev. Clin. Oncol. 12, 511–526
393   (2015).
394   22.   Hammerman, P. S. et al. Comprehensive genomic characterization of squamous
395   cell lung cancers. Nature 489, 519–525 (2012).
396   23.   Collisson, E. A. et al. Comprehensive molecular profiling of lung adenocarcinoma:
397   The cancer genome atlas research network. Nature 511, 543–550 (2014).
398   24.   Keren, L. et al. A Structured Tumor-Immune Microenvironment in Triple Negative
399   Breast Cancer Revealed by Multiplexed Ion Beam Imaging. Cell 174, 1373-1387.e19 (2018).
400   25.   Giesen, C. et al. Highly multiplexed imaging of tumor tissues with subcellular
401   resolution by mass cytometry. Nat. Methods 11, 417–422 (2014).
402   26.   Goltsev, Y. et al. Deep Profiling of Mouse Splenic Architecture with CODEX
403   Multiplexed Imaging. Cell 174, 968-981.e15 (2018).
404
405

450 **Author Contributions**

464 **Competing Interests**

479 **Materials and Correspondence**

480 Materials request and general correspondence should be addressed to J.L.Q., C.S. and Y.Y.

481

482 **Data availability**

483 The digital pathology images from the TRACERx study generated or analysed during this
484 study are not publicly available and restrictions apply to its use. A test subset of such

485   digital pathology images are available through the Cancer Research UK & University
486   College London Cancer Trials Centre (ctc.tracerx@ucl.ac.uk) for non-commercial research
487   purposes and access will be granted upon review of a project proposal that will be
488   evaluated by a TRACERx data access committee and entering into an appropriate data
489   access agreement, subject to any applicable ethical approvals. Digital pathology images for
490   LATTICe-A samples with expert pathologist's annotations used for validation are available:
491   https://github.com/qalid7/compath. Request for data access for the remaining LATTICe-A
492   samples can be submitted to J.L.Q.
493

494   **Code availability**
495   The deep learning pipeline for digital pathology image analysis is available for
496   noncommercial research purposes: https://github.com/qalid7/compath. All code used for
497   statistical analyses of image data was developed in R version (3.5.1) and is available:
498   https://github.com/qalid7/tx100_compath.

499

Methods

502

Tissues and digital images

The main cohort evaluated comes from the first 100 patients prospectively analyzed by the lung TRACERx study[6] (Extended Data Fig. 1, Supplementary Tables 1, 4, https://clinicaltrials.gov/ct2/show/NCT01888601, approved by an independent Research Ethics Committee, 13/LO/1546). 62 were men and 38 were women, with a median age of 68. 61 were LUAD, 32 were LUSC and the remaining 7 had 'other' histology subtypes (including adenosquamous carcinoma, large cell carcinoma, large cell neuroendocrine carcinoma, pleomorphic carcinoma and pleomorphic carcinoma arising from adenocarcinoma).

512

The 85 case subcohort with regional histology consisted of 55 male and 30 female patients and of those 49 were LUAD, 32 were LUSC and 6 were 'other' types. 10 of these patients had a single region while the rest ranged between 2-8 regions (= 275 total regional histology samples). Snap-frozen regional samples were processed to FFPE blocks after dissecting fresh-frozen tissues for DNA-seq and RNA-seq analyses. Tissue microarrays (TMAs) were created containing 133x2mm regional tissue cores from 75 patients in 7 blocks.

520

In addition to the regional samples, full-sized diagnostic blocks were obtained for all 100 cases precisely mirroring the Jamal-Hanjani et al. 2017 prospective 100 patient cohort[6]. 4μm thick sections were cut and subjected to H&E staining and multiplex IHC for CD8/CD4/FOXP3: anti-CD8 (type: Rabbit Monoclonal, clone: SP239, cat. no.: ab178089, source: Abcam Plc, Cambridge, UK, used at 1:100); anti-CD4 (type: Rabbit Monoclonal, clone: SP35, cat. no.: ab213215, source: Abcam Plc, Cambridge, UK, used at 1:50); antiFOXP3 (type: Mouse, clone: 236A/E7, source: kind gift from Dr G Roncador, CNIO, Madrid, Spain, used at: 1:100). All regional and diagnostic slides were scanned using NanoZoomer S210 digital slide scanner (C13239-01) and NanoZoomer digital pathology system version 3.1.7 (Hamamatsu, Japan) at 40x (228 nm/pixel resolution).

532

The external validation cohort was obtained from the Leicester Archival Thoracic Tumor Investigatory Cohort – Adenocarcinoma (LATTICe-A) study[8], a continuous retrospective series of resected primary LUAD tumors from a single surgical center between years 1998 to 2014 (Extended Data Fig. 1, Supplementary Table 5). It consists of 4,324 whole-tumor diagnostic blocks from 970 LUAD patients (ranging from 1 to 16 blocks per case with a median of 4). 455 were men and 515 were women with a median age of 69. Most clinical data (age, sex, adjuvant therapy status and time to recurrence or death) were available for

540 all patients, with complete pathological stage for 827 and smoking history for 651. All
541 archival slides containing tumor material were used in order to capture the full diversity of
542 each lesion. Slides were dearchived and scanned using a Hamamatsu NanoZoomer XR at
543 40x (226 nm/pixel resolution) yielding 15 TB of image data. Images containing incidental
544 lymph node tissue were excluded to avoid confounding immune infiltration analysis. For
545 the biological validation assay, a subset of 49 paraffin blocks from 49 patients was
546 obtained from the same study, and from these a validation TMA was prepared, containing
547 a single 1mm core from each case. The work was ethically approved by an NHS research
548 ethics committee (ref. 14/EM/1159). This study complies with the STROBE guidelines.
549

550 **The deep learning pipeline for cell detection and classification**
551 The deep learning pipeline consists of three parts. First, the pipeline segments tissue
552 regions utilizing multi-resolution input/output image features (Micro-Net[27]). It was
553 designed to capture global tissue context and learn weak features that could be important
554 for identifying tissue boundary, but are often not achieved by other machine learning
555 methods such as thresholding of the grey-scale image, active contours, watershed
556 segmentation or Support Vector Machine-based training on local binary pattern features[27].
557 Tissue segmentation removes background noise and artefacts and subsequently allows for
558 more computationally efficient cell detection and accurate classification. Secondly, a cell
559 detection model modified from SCCNN[28] predicts for each pixel the probability that it
560 belongs to the center of a nucleus within tissue regions identified by Micro-Net. Nuclei are
561 detected from the probability map obtained from the deep network. Lastly, a cell
562 classification framework utilizes a neighboring ensemble predictor classifier coupled with
563 SCCNN to classify each cell by type.
564

565 For tissue segmentation, each whole slide image was reduced to 1.25x resolution and
566 segmented for tissue regions using Micro-Net-512[27] architecture. This architecture
567 visualizes the image at multiple resolutions, captures context information by connecting
568 intermediate deep layers and adds bypass connections to max-pooling to maintain weak
569 features (Fig. 1b). 10 whole slide images were used to train the tissue segmentation
570 network using MicroNet. The segmented images from the network were inspected visually
571 and quantitatively (Supplementary Table 6, Supplementary Figures 1-20) to evaluate
572 performance using an independent set of images.
573

574 The SCCNN adds two layers to conventional deep learning architecture for cell detection
575 within the segmented tissue. SC1 estimates the location and probability of each pixel
576 belonging to the center of a cell, and these probabilities are then mapped by SC2 to the
577 image. A customized implementation of SCCNN was coded in Python (version 3.5) using
578 TensorFlow[29] library (version 1.3) which makes it computationally more efficient
579 compared to the original MATLAB implementation[28]. To process an image of size

580 1000×1000 pixels, the Python implementation takes 4.8 seconds for nucleus detection
581 compared to 41.0 seconds using the original implementation[28], excluding preprocessing
582 which remained the same in both implementations (using MATLAB (version 2018b)). In
583 addition, through empirical experimentation, we optimized the patch size to 31x31 instead
584 of 27x27 in the original implementation for increased cell detection accuracy. To generate
585 nuclear locations from the SC2 probability map, peak detection was applied where
586 thresholds for intensity and minimum grouping distance were also optimized to 0.15 and
587 12 pixels through experimentation using validation data.
588
589 For cell classification, a neighboring ensemble predictor was used. This predictor utilizes
590 SCCNN to classify cells in neighboring locations to the detected center of the cell. In our
591 implementation, the ensemble classifier required votes from SCCNN classification of nine
592 different neighborhood locations near to the center of the cell compared to five votes in
593 original implementation. Through experimentation, the patch size was optimized to 51x51
594 for classification instead of 27x27 as originally proposed. This permitted incorporation of
595 greater tissue spatial context while maintaining the accuracy of classifying small cells.
596
597 Altogether, this pipeline enabled the spatial mapping of four cell types from H&E images:
598 cancer (malignant epithelial) cells, lymphocytes (including plasma cells), non-inflammatory
599 stromal cells (fibroblasts and endothelial cells), and an "other" cell type that included
600 nonidentifiable cells, less abundant cells such as macrophages and chondrocytes, and
601 'normal' pneumocytes and bronchial epithelial cells.
602

603 **Training the deep learning pipeline**

604 To improve neural network generalizability and to avoid overfitting for cell detection and
605 classification, we trained and tested our pipeline on a variety of sample types, including
606 diagnostic (= 100), regional (= 275) and 133 cores corresponding to 75 TRACERx patients
607 from TMA slides (63 patients had two cores and 12 patients had a single core). Both cell
608 detection and classification were trained based on single-cell annotations from
609 pathologists. Two thoracic pathologists annotated 26,960 cells on 53 whole slide images (3
610 TMAs, 35 regional slides and 15 diagnostic slides) to incorporate morphological variations
611 in appearance of various cell types and stain variability. Several hundred examples of each
612 cell class were marked on 76 cores selected at random from TMA images. In total, 4,056,
613 5,310, 15,007, 2,587 annotations were collected for stromal cells, lymphocytes, cancer
614 cells and "other" cell types, respectively. These whole slide images were divided into small
615 tile images of size 2000×2000 pixels (each pixel = 0.5µm), which were then divided into
616 three sample sets maintaining the class distribution of cells. These included: 13 diagnostic,
617 58 regional and 134 TMA tile images for training; 4 diagnostic, 21 regional and 72 TMA tile
618 images for validation; and 3 diagnostic, 22 regional and 61 TMA tile images for testing. As

619   a result, the annotations were divided between the three groups; 2/3 for training, 1/6 for

620   validation and 1/6 for testing. The training set included annotations for 2,147 stromal cells,

621   3,183 lymphocytes, 10,103 cancer and 1,357 other cell types. The validation set had

622   annotations for 473 stromal cells, 825 lymphocytes, 2,562 tumor and 359 other cell types.

623   Breakdown for the test set is provided in Supplementary Table 2.

624

625   For IHC cell classification, we used a pretrained SCCNN network on samples stained for

626   CD4/CD8/FOXP3. The training set consisted of 1,657 $CD4^+FOXP3^-$, 3,187 $CD8^+$, 1,001

627     $CD4^+FOXP3^+$, and 3,488 other (negative) cells. The trained network was tested on 5,028

628   cell annotations collected on 6 lung diagnostic whole slide images, including 251

629     $CD4^+FOXP3^-$, 406 $CD8^+$, 123 $CD4^+FOXP3^+$ and 4,248 other cells to test the ability of the

630   algorithm in correctly detecting and classifying negative cells. See Supplementary Table 7

631   for the total number of identified cells in the H&E diagnostic, H&E multi-region and IHC

632   diagnostic datasets.

633

634   **Validation of the H&E deep learning pipeline with orthogonal data types**

635   The algorithms' performance in detecting and classifying single cells in H&E were first

636   evaluated against the test set of 5951 cells. Individual class accuracy statistics were

637   calculated using the R function 'confusionMatrix' from the R package 'caret'.

638

639   Pathology TIL estimates were scored following the international guidelines developed by

640   the International Immuno-Oncology Biomarker Working Group[7]. Briefly, by inspection of

641   H&E slide of a given tumor region, the fraction of the stromal area infiltrated by TILs was

642   assessed.

643

644   For regional samples, tumor cellularity, estimated as the computed percentage cancer cells

645   was correlated with tumor purity estimated by ASCAT based on DNA-seq copy number and

646   VAF purity (both available from Jamal-Hanjani et al.[6], = 239 regional tumor samples). The

647   RNA-seq-based $CD8^+$ T cell signature (available from Rosenthal et al.[1], computed using the

648   Danaher et al. method[30]) was correlated with the deep learning based lymphocyte

649   percentage for 142 regional tumor samples. For diagnostic samples, deep learning-based

650   lymphocyte percentage from H&E was correlated with deep learning-based $CD8^+$ cell

651   percentage from IHC (= 100 diagnostic samples, Extended Data Fig. 2a-d).

652

653   Discordance rate between RNA-seq based[1] and histology/deep learning-based immune hot

654   and cold regional classification was calculated by cross-tabulation of immune hot and cold

655   (from histology) versus high and low (from RNA-seq), disregarding any regions without one

656   of these two types of data. The RNA-seq method used 15 immune cell signatures

657   presenting different T- and B-cell subsets, as well as neutrophils, macrophages, mast and

dendritic cells, to classify tumor regions into high and low categories. A Fisher's exact test was used to compute the overlap between the two immune classifications. Distributions of multiple immune scores (lymphocyte percentage, intra-tumor lymphocytes and adjacent-tumor lymphocytes/stroma) as well as ASCAT tumor purity were compared between hot versus cold (deep learning) and high versus low (RNA-seq) classifications (Extended Data Fig. 5).

Validation of the deep learning pipeline with the independent LATTICe-A cohort

The external validity of the proposed deep learning pipeline was performed on 100 randomly selected patients from the LATTICe-A cohort[8]. This validation ensures that the trained cell detection and cell classification models from the TRACERx tumor blocks are generalizable to a distinct dataset which is processed, stained and scanned in another center (the LATTICe-A study, University of Leicester).

All 100 whole-tumor H&E sections were processed using the same TRACERx trained model. The validation was then performed using two data types. First, a pathologist provided 5,082 single-cell annotations following the same protocol for TRACERx in 20 randomly selected LATTICe-A sections. The breakdown for single-cell annotations was 1,997 stromal cells, 787 lymphocyte cells, 1,839 cancer cells and 459 other cells (see Supplementary Table 3). Second, two independent pathologists jointly scored the remaining 80 sections for overall fraction of lymphocytic infiltration and pathology TIL estimates[7]. These manual scores were correlated with the deep learning-based lymphocyte percentage and adjacent-tumor lymphocytes/total stroma (Extended Data Fig. 2e).

Validation of the deep learning pipeline with biological assays

A new biological validation method was developed to overcome the challenge of obtaining large quantities of cell-specific validation data (Fig. 1f-h, Extended Data Fig. 2f-g). 48 cores were available for the TTF1-H&E image pairs, 38 for the CD45-H&E pairs, and 33 for the SMA-H&E pairs. Stains were performed using a Ventana BenchMark ULTRA instrument (H&E, TTF-1) or a Dako Link 48 (CD-45, SMA). Digital images were acquired using a Hamamatsu Nanozoomer slide scanner. First, H&E staining was performed using a Leica Infinity kit, and a digital image was collected. The slide was subsequently de-coverslipped, the H&E stain removed by acid alcohol washing, and then an immunohistochemical stain with haematoxylin counterstain was applied using a standard diagnostic antigen retrieval and antibody protocol. A second digital image was acquired after mounting and coverslipping. Through experimentation, no difference in the staining was observed when the procedure was reversed.

696 TTF-1 (type: Novocastra Liquid Mouse Monoclonal antibody thyroid transcription factor 1,
697 clone: SPT24, cat. no.: NCL-L-TTF-1, source: Leica biosystems, Germany, used at 1:100) was
698 selected as the cancer cell marker in these LUAD samples because it is the most robust and
699 widely used immunohistochemical marker of LUAD cells[31]. It is very specific, both in that
700 only epithelial cells are stained in the lung, and in that very few tumors of non-lung or
701 thyroid origin are stained[32]. The sensitivity of the antibody clone used (SPT24) is also high,
702 staining >75% of tumor cells in 76% of LUAD tumors in one published series[33]. However, as
703 this implies, there are many tumors in which tumor cell staining is incomplete (i.e. <100%).
704 Therefore, only cores showing near-universal TTF-1-positivity of tumor cells were used for
705 validation, in order to provide the best possible 'gold standard' comparator for the deep
706 learning algorithm. The same procedure was followed for pairs of H&E-CD45 (anti-human
707 CD45, type: Mouse Monoclonal, clone: 2B11 + PD7/26, cat. no.: M0701, source: Agilent
708 DAKO, USA, used at 1:200) and H&E-SMA (myofibroblast marker, type: Mouse Monoclonal
709 antibody Smooth Muscle Actin (1A4), cat. no.: 760-2833, source: Roche, Switzerland, a
710 ready to use antibody) to biologically validate the accuracy of single cell classification.

711

712 In total, 64,976 TTF1$^+$ cells, 26,284 CD45$^+$ cells and 46,343 SMA$^+$ cells were detected from
713 the IHC images, denoting the advantage of this method in acquiring large amount of
714 validation data at single-cell resolution. The correlation measured (Fig. 1f-h, Extended Data
715 Fig. 2g) was that between the fraction of classified cells in the H&E versus fraction of
716 positively stained IHC cells per 100μm^2.

717

718 **Immune phenotype classification**
719 To classify tumor regions into different immune phenotypes, we assigned each region to
720 an immune hot, cold or intermediate category based on lymphocyte percentage. The
721 dependency of our subsequently results on thresholds chosen for this classification
722 scheme was tested after applying perturbations to the thresholds used. Four new
723 classification schemes were tested: no intermediate zone (i.e. using median lymphocyte
724 percentage for separating hot and cold regions), regions with lymphocyte percentage
725 greater than standard deviation/2 above/below the median lymphocyte percentage
726 classified as immune hot/cold, , and similarly for standard deviation/3 and standard
727 deviation/6 (Extended Data Fig. 4a-b). For every new classification, we repeated the
728 multivariate survival analysis to confirm the significance of the number of immune cold
729 regions in predicting disease-free survival as well as the genomic distance test for pairs of
730 immune hot versus immune cold regions in LUAD patients (Extended Data Fig. 4b). In
731 addition, the CD8$^+$ RNA-seq signature was used to test the difference in CD8$^+$ levels
732 between immune hot and immune cold phenotypes across all classification schemes
733 (Extended Data Fig. 4c).

734

Genomic distance measure

Genomic distance was calculated as described previously[1], by taking the Euclidean distance of the mutations present for every pair of immune hot and immune cold regions from the same patient. All mutations present in a region from a tumor were turned into a binary matrix of which the rows were mutations and columns were the tumor regions. From this matrix, the pairwise distance was determined.

Distance between dominant clones to the last common ancestor of region pair

Deep learning-based immune phenotypes were integrated with the TRACERx phylogenetics data[6]. Dominant clones (using the upper quartile of cancer cell fraction, ≥ 75%) were labelled for all tumor regions' trees which had an available H&E sample in LUAD patients (= 76 regions, 15 immune hot pairs and 23 immune cold pairs). For every pair of immune hot / cold regions within a tumor, the distance between the dominant clones (as measured by branch length, i.e. number of mutations) via their last common ancestor was computed. The recently shared ancestry clone between the two dominant clones was labelled as the 'last common ancestor of region pair' (annotated with arrows in Fig 3.c). To ensure this analysis was not dependent on a certain cancer cell fraction threshold, multiple thresholds (CCF ≥ 80%, 85%) were placed while repeating the same analysis. Next, by identifying the last common ancestral subclone for pairs of the same phenotype, each pair was categorized into one of two diversification patterns: 'diversifying at the most recent common ancestor (MRCA) of the tree' or 'diversifying at a descendant subclone of the MRCA of the tree'. The latter category included a pattern exclusive to immune cold pairs, where the two regions shared the same dominant subclone that was the direct descendant of the MRCA of the tree.

Tumor spatial modelling

H&E and IHC cell abundance scores (e.g. lymphocyte percentage, CD8$^+$ percentage) were computed as the percentage of a cell type in the total sample cell count. Stromal TILs were identified using spatial modelling[20,34,35], where lymphocytes were classified (using unsupervised clustering) into intra-tumor lymphocytes, adjacent-tumor lymphocytes and distal-tumor lymphocytes based on their spatial proximity to epithelial cell nests in H&Es. The immune hotspot score was calculated using the Getis–Ord algorithm as previously described[36]. To capture the emergence of complex morphological patterns that dictate cancer-stromal cell spatial contact preserved over varying spatial scales, a fractal dimension calculation (Minkowski-Bouligand dimension) was performed using the box-counting algorithm[37]. This algorithm calculates the number of boxes of a certain size needed to cover a geometric pattern. We modified a MATLAB-based algorithm[38] to include both spatial information of cancer and stromal cells, as opposed to its conventional use on one variable (i.e. pixel information of an image). The analysis was carried out on spatial

774   maps generated using coordinates of classified stromal and cancer cells, while utilizing the

775   tissue segmented image (as a boundary mask) to exclude all empty tissue areas. Choices of

776   box size were informed by the distribution of minimum and maximum Euclidean distance

777   for each stromal cell to its nearest cancer cell in all 275 tumor regions (Extended Data Fig.

778   7a). The mean minimum distance was 21.43μm. We limited the upper box size at 300μm,

779   which is just above a previously proposed cell-cell communication distance of 250μm[39] but

780   designed to be more inclusive. For statistical tests where fractal dimension was

781   represented at tumor level, the maximum regional score was used.

782

783   **H&E-IHC spatial alignment/immune subset projection**

784   For a H&E diagnostic slide, we determined the number of intra-tumor lymphocytes,

785   adjacent-tumor lymphocytes and distal-tumor lymphocytes (n_I, n_A, n_D) based on spatial

786   modelling of the H&Es. After spatial alignment of IHC and projecting IHC-derived cells onto

787   the H&E, the number of $CD8^+$ cells that were also intra-tumor lymphocytes was

788   determined (n^{CD8}_{ITL}), and similarly for other cell types. As a result, intra-tumor lymphocytes

789   were deconvoluted by $n_I = n^{CD8}_I + n^{CD4}_I + n^{FOXP3}_I + n^{other}_I$. Two-sided paired Wilcox was used to test

790   the difference in the percentage of $CD8^+$ cells among intra-tumor lymphocytes, adjacent-

791   tumor lymphocytes and distal-tumor lymphocytes (n^{CD8}_{ATL}, n^{CD8}_{DTL}, n^{CD8}_{ITL}). The same test was

792   performed for $CD4^+FOXP3^-$ and $CD4^+FOXP3^+$ cells.

793

794   The 10 LUAD patients with the highest adjacent-tumor lymphocytes to stromal cell ratio

795   were selected for this immune subset spatial projection. All samples had above median

796    $CD8^+$%. One sample was excluded due to poor HE-IHC alignment quality and the

797   subsequent analysis was performed on the remaining nine samples. The quality of

798   alignment was evaluated by manually identifying 238 visible landmarks and placed on

799   corresponding positions in H&E and IHC tiles (total number of tiles = 249, maximum

800   landmarks per tile = 5), as shown in Extended Data Fig. 8b. These marked points were used

801   to compute the Euclidean distance (difference in , coordinates) between them to obtain a

802   quantitative measurement of alignment accuracy. The average distance between matching

803   landmarks was 9.57μm, whereas the maximum distance between the H&E and

804   CD4/CD8/FOXP3 sections was 16μm.

805

806   **Survival analysis and other statistical methods**

807   Survival tests were conducted using Kaplan-Meier estimator ('ggsurvplot' R function from

808   the 'survminer' and 'survival' R packages) as well as Cox model ('coxph' R function and

809   displayed using 'ggforest' R function). Forest plots show the hazard ratio in the x-axis; each

810   variable's hazard ratio is plotted and annotated with a 95% confidence interval. The

811   clinical parameters included in the multivariate model were age, sex, smoking pack years,

812   histology (whether LUAD, LUSC or otherwise), tumor stage, adjuvant therapy (whether

813   received or not). Because of its prognostic importance in TRACERx, the upper quartile of
814   clonal neoantigens in each histology cohort was also incorporated in the multivariate
815   model. The range of available disease-free survival data was 34-1364 days (median = 915
816   days) in TRACERx, and 1-6139 days (median = 684 days) in LATTICe-A. All hazard ratios
817   were computed on all time points (i.e. the whole survival curve, not at a specific time
818   point). Correlation tests used Spearman's method and were generated using the function
819   'ggscatter' from the 'ggpubr' R package. All correlation plots show the Rho (ρ) coefficient
820   and the significance R-value. For statistical comparisons among groups, a two-sided,
821   nonparametric, unpaired, Wilcoxon signed-rank test was used, unless stated otherwise. All
822   box plots were generated using the function 'ggboxplot' from the 'ggpubr' R package (all
823   data points are plotted with the 'jitter' option, the median value is indicated by a thick
824   horizontal line; minimum and maximum values are indicated by the extreme points; the
825   first and third quantiles are represented by the box edges; and vertical lines indicate the
826   error range) or the function 'ggbetweenstats' from the 'ggstatplot' R package for more
827   than two groups. Tests for concordance between two data classes were analyzed using a
828   Fisher's exact test. All statistical tests were two-sided, a R value of less than .05 was
829   considered statistically significant. To adjust R-values for multiple comparisons, the
830   Benjamini & Hochberg method was used. To measure effect size, Cohen's d method was
831   used. All statistical analyses were conducted in R (version 3.5.1).
832
833   **Reporting summary**
834   Further information on research design is available in the Nature Research Reporting
835   Summary linked to this paper.
836

Extended Data Figures legends

Extended Data Fig. 1. CONSORT diagrams for TRACERx 100 and LATTICe-A histology cohorts and patient characteristics. a. TRACERx CONSORT diagram to illustrate sample collection and analysis of regional and diagnostic histology samples, as well as the overlap with RNA and DNA studies. **b.** TRACERx patient characteristics for the histology cohort. **c.** LATTICe-A CONSORT diagram (= 970 LUAD patients). Legends for 'type of the analysis' correspond to panel **a**. **d**. Demographics and clinical patient characteristics for TRACERx (top three panels) and LATTICe-A (bottom three panels) showing the distribution of age (colored by sex), distribution of smoking pack years and the proportion of patients in each pathological stage. Horizontal lines indicate the median value.

Extended Data Fig. 2. Validation of the automated single-cell classification for H&E. a. A scatter plot showing the correlation between H&E-based adjacent-tumor lymphocytes/stromal and pathology TIL estimates in diagnostic samples (= 98 diagnostic slides/patients). **b.** Scatter plots showing the correlations between H&E-based tumor cellularity estimate and ASCAT/VAF purity scores (= 238 regions; 83 patients). **c.** A scatter plot showing the correlation between H&E-based estimate of lymphocyte percentage among all cells and RNA-seq-based CD8[+] signature using the Danaher et al. method[30] (= 142 regions; 56 patients). **d.** A scatter plot showing the correlation between H&E-based estimate of lymphocyte percentage among all cells and CD8[+] cell percentage in IHC in the diagnostic samples (= 100 diagnostic slide/patients). **e.** Scatter plots showing the correlation between H&E-based lymphocyte percentage versus pathological scores of overall lymphocytic cell fraction, and adjacent-tumor lymphocytes/stromal versus pathology TIL estimates in an external cohort (LATTICe-A, = 80 diagnostic slides/patients). **f.** Illustrative example to show the spatial alignment of TTF1/CD45/SMA-stained IHC and H&E images obtained using sequential staining on the same tissue microarray section for biological validation. **g.** A scatter plot showing the correlation between stromal cell percentage determined by H&E and SMA[+] cell percentage per LUAD image tiles of size $100\mu m^2$ (= 144). The experiment was conducted once using one TMA (= 33 cores/patients). The shading indicates 95% confidence interval.

Extended Data Fig. 3. Distribution of regional lymphocytic infiltration according to pathological stage. All available patients' data have been used in this figure except for the standard deviation tests excluding patients with a single tumor region. Patients without pathological staging information from the LATTICe-A cohort were also removed. **a, b, c**, top row: TRACERx and bottom row: LATTICe-A. Horizontal lines indicate the median value.

a. Distribution of the standard deviation of regional lymphocyte percentage for LUAD and LUSC patients in TRACERx (= 69), and LUAD in LATTICe-A (= 814). **b.** Distribution of the standard deviation of regional lymphocyte percentage across pathological stages (= 69 for TRACERx, 814 for LATTICe-A). **c.** Distribution of regional mean of lymphocyte percentage across stages (= 79 for TRACERx, 827 for LATTICe-A). **d**. No significant difference among stages with respect to standard deviation (= 69 for TRACERx, 814 for LATTICe-A) or mean (= 79 for TRACERx, 827 for LATTICe-A) of regional lymphocytic infiltration. Left panel, TRACERx and right panel, LATTICe-A. Correction for multiple testing was applied in **d**, for each cohort individually. A two-sided, non-parametric, unpaired, Wilcoxon signed-rank test was used; each dot represents a patient; the mean value is annotated with a large dot; the median value is represented by a thick horizontal line; minimum and maximum values are indicated by the extreme points; the first and third quantiles are represented by the box edges; and the violin shape shows the data distribution as a kernel density estimation.

Extended Data Fig. 4. Validation of immune phenotype classification. a. The proposed immune classification imposed on density plot showing distribution of lymphocyte percentage. The middle zone corresponds to the intermediate phenotype, red zone for immune hot and blue zone for immune cold. Black dash line shows the median. This classification was validated after applying small perturbations to the thresholds to re-classify regional immune phenotypes, illustrated as grey dash lines: no intermediate zone (i.e. hard median for separating hot and cold), standard deviation (SD)/2 above and below the median, SD/3 and SD/6. **b**. Forest plots to show repeated multivariate Cox regression tests for the number of immune cold regions using these new classifications (= 79 patients), after accounting for stage, total number of samples, upper quartile of clonal neoantigens determined for LUAD and LUSC individually, and other clinical parameters. Box plots showing difference in genomic distance for pairs of hot regions compared with pairs of cold regions for LUAD and LUSC separately (LUAD: = 45 hot pairs, 45 cold pairs for no intermediate zone; = 19 hot, 25 cold for SD/2; = 25 hot, 33 cold for SD/3; = 32 hot, 41 cold for SD/6. LUSC: = 32 hot pairs, 54 cold pairs for no intermediate zone; = 19 hot, 27 cold for SD/2; = 19 hot, 37 cold for SD/3; = 27 hot, 41 cold for SD/6.). **c**. Box plots showing significant difference in $CD8^+$ RNA-seq signature using the Danaher method between regions of hot and cold phenotype across all classification schemes (= 219 for SD/4; 275 for no intermediate zone; 173 for SD/2; 204 for SD/3; 237 for SD/6). **d**. Distribution and difference of lymphocytic infiltration for LUAD versus LUSC regions in TRACERx (= 275 regions; 85 patients) as well as distribution for LUAD in LATTICe-A (= 4,324 samples; 970 patients). Horizontal lines in the distribution plots indicate mean values. For statistical comparisons among groups, a two-sided, non-parametric, unpaired, Wilcoxon signed-rank test was used, unless stated otherwise.

914

Extended Data Fig. 5. Concordance between histology deep learning and RNA-seq immune classification. a. A box plot showing the difference in pathology TIL estimates between immune hot and immune cold regions (= 219). Pathology TIL estimates score fraction of stroma containing TILs, whereas immune classification was defined based on the percentage of lymphocytes in all cells within a slide. **b**. A confusion matrix to compare RNA-seq and deep learning histology immune classifications (discarding immune intermediate regions, = 109 regions (57 LUAD, 37 LUSC, 15 other histology subtypes); 52 patients). The p-value was generated using a two-sided Fisher's exact test for overlap. **c**. A box plot showing the difference in the fraction of immune hotspots[36] in regions where the two classifications are in agreement (= 78; labeled as 'In agreement') against the discrepant regions (= 31, labeled as 'Discrepant'). Each dot represents a region, the median value is indicated by a thick horizontal line; minimum and maximum values are indicated by the extreme points; and the first and third quantiles are represented by the box edges. **d**. Box plots to support the overall consistency between H&E-deep learning and RNA-seq methods by comparing different immune scores as well as ASCAT tumor purity between immune hot/high and cold/low tumor regions (all R -values < 0.0001). Top row, H&E-deep learning immune classification (= 219; except the ASCAT purity box plot = 186 regions), bottom row, RNAseq derived immune classification (= 142; except the ASCAT purity box plot, = 141 regions). For statistical comparisons among groups, a two-sided, non-parametric, unpaired, Wilcoxon signed-rank test was used, unless stated otherwise.

935

Extended Data Fig. 6. Genomic and survival analysis of tumor regions according to immune phenotypes. a. A box plot showing the difference in genomic distances for pairs of immune hot versus immune cold regions within the same LUSC patients (= 59 pairs). A two-sided, non-parametric, unpaired, Wilcoxon signed-rank test was used. **b.** Forest plots to show the univariate prognostic value for the number of immune low regions (both as continuous and dichotomized at the median (≤1 versus >1)), or the number of immune high regions, using the immune classification generated by RNA-seq-based infiltrating immune cell populations[1] in 64 TRACERx tumors (41 LUAD, 16 LUSC and 7 other histology subtypes)**.**
c. Forest plots showing multivariate Cox regression analyses in both TRACERx (= 79 patients; LUAD and LUSC combined) and LATTICe-A (= 651 LUAD patients representing a subset with complete stage and smoking pack years data) with the number of immune cold regions dichotomized at the median (≤1 versus >1). This remains significant when the number of immune cold regions was replaced as a continuous variable, in the same multivariate model, (R = 0.019 in TRACERx and < 0.001 in LATTICe-A, for the number of immune cold regions). Clonal neoantigens were dichotomized using the upper quartile,

952 determined individually for LUAD and LUSC tumors[1]. **d**. The same test in **c** when tumor size
953 (in mm) was also controlled in the multivariate model in LATTICe-A. This test also
954 remained significant for a bigger group of patients with complete stage data, but missing
955 pack years information (= 815, R < 0.001, HR = 1.4[1.1-1.8]). **e**. Forest plots to compare
956 the prognostic value of regional immune scores as well as diagnostic H&E and IHC scores
957 for relapse-free survival in TRACERx (= 79 patients, LUAD and LUSC combined). Wherever
958 possible, these immune features were tested in LATTICe-A (= 970 patients). To compare
959 the prognostic value of the number of immune cold region with other immune features,
960 LATTICe-A comparisons were conducted in Cox multivariate regression models to include
961 every immune feature after correcting for the number of immune cold regions in the same
962 model. Each variable's HR is plotted with a 95% confidence interval; all R-values were
963 adjusted for multiple testing; and the size of the circles denotes $-\log_{10}(R$).   For   the
964 sake of visualization, a minor adjustment was made to the HR for the number of cold
965 regions/total number of regions in LATTICe-A from 0.88[0.57-1.3] to 0.99[0.97-1.3]. SD:
966 standard deviation, used for measuring variability of lymphocyte percentage among
967 samples within a tumor. **f.** Forest plots using Cox multivariate regression analysis showing
968 that the prognostic value of the number of immune cold regions was independent of: 1)
969 genetic measure, subclonal copy number alteration (obtained from [6]); 2) tumor cellularity
970 from DNA-seq-based ASCAT purity, 3) tumor cellularity measured by deep learning-based
971 cancer cell percentage. **g.** Kaplan Meier curves to illustrate the difference in relapse-free
972 survival for TRACERx patients including other histology types (= 85; representing all
973 TRACERx patients in the multiregion histology cohort) with high and low number of
974 immune cold regions, dichotomized by its median value. Log-rank R = 0.0017. **h.** Forest
975 plot using Cox regression for the multivariate survival analysis for the number of immune
976 cold regions in TRACERx including patients with other histology subtypes (= 85).

977

978 **Extended Data Fig. 7. Fractal dimension and relationships with stromal cells. a.**
979 Distribution of the average minimum Euclidean distance between a stromal cell to its
980 neighboring cancer cell. For every stromal cell in a tumor region slide, the minimum
981 distance to nearest cancer cell was computed. This distance was then averaged for all
982 identified stromal cells in every region to plot the distribution (= 275 regions; 85 patients).
983 **b**. Distribution of the fractal dimension of the cancer-stroma cell interface for histology
984 types in the TRACERx cohort (= 275 regions; 85 patients). **c**. Box plots to show the
985 difference in fractal dimension between immune hot and cold regions in TRACERx LUAD (=
986 113) and LUSC (= 84). **d**. Box plots showing the difference in stromal cell percentage
987 between immune hot and cold regions in all (= 219), LUAD (= 113), and LUSC (= 84). **e.**
988 Scatter plots showing the correlation between fractal dimension and percentage of cells
989 that are stromal or cancer in all tumor regions (= 275 regions; 85 patients). This shows
990 that fractal dimension was independent of tumor cell composition, with only a weak

correlation with stromal cell percentage and no correlation with tumor cellularity. **f**. Box
992 plots showing the difference in fractal dimension between LUAD tumor regions harboring
993 an LOH event for HLA type A (= 106), type B (= 113), type C (= 108) versus regions that do
994 not, adjusted for multiple comparisons with the corresponding test in Fig. 4c. **g**. The same
995 test in **f** repeated for LUSC tumor regions (= 87) for HLA of any type. **h.** Box plots showing
996 the difference in tumor-level fractal dimension using the maximum value of regional
997 measures between LUAD tumors (= 48) harboring a single LOH event for any HLA type,
998 HLA type A, type B and type C versus tumors that do not, independent of predicted clonal
999 neoantigens. Each p-value was generated using a multiple regression linear model and was
1000 also adjusted for multiple testing correction. **i**. The same test in **h** repeated for LUSC
1001 tumors (= 29) for HLA of any type. For statistical comparisons among groups, a two-sided,
1002 non-parametric, unpaired, Wilcoxon signed-rank test was used, unless stated otherwise.

1003

1004 **Extended Data Fig. 8. Relationship of immune subsets and spatial TILs in LUAD. a.**
1005 Spearman's correlations between immune scores in diagnostic slides and genetic
1006 measures including predicted neoantigens and HLALOH in LUAD patients (= 46). ITLR:
1007 intra-tumor lymphocytes to total tumor cell ratio. Only significant correlations after
1008 multiple testing are highlighted (rho = 0.37, R = 0.035). **b**. Examples of registered H&E and
1009 IHC tiles. The green cross denotes a manually placed landmark repeated 238 times on
1010 pairs of H&E-IHC image tiles. The Euclidean distance (difference in , coordinates) was
1011 computed between the two landmarks which was then **c.** shown as a distribution to
1012 represent the accuracy of the registration (= 249 total H&E-IHC image tiles, maximum five
1013 landmarks per a pair of tiles). The average distance between matching landmarks was
1014 9.57μm and the distribution is within the expected range of maximum distance between
1015 four serial sections (16μm). **d.** Box plots to illustrate the difference in percentage of
1016 immune cell subsets among adjacent, intra and distal-tumor lymphocytes (= 20 image
1017 tiles), a non-parametric, paired Wilcoxon test was used.

1018

1019 **Extended Data Fig. 9. Summary of immune and genomics features in NSCLC.** An extended
1020 heatmap showing all immune variables described in TRACERx across all patients (= 275
1021 regions; 85 patients), along with genetic measures and clinical parameters. Each column
1022 represents a tumor, grouped by their histologic subtype. Tumor regions (illustrated as
1023 dots) were assigned to immune hot, immune cold and intermediate phenotypes based on
1024 percentage of lymphocytes in all cells following H&E-based deep learning analysis.
1025 Cancerstromal fractal dimension, defined using the maximum fractal dimension in regions
1026 of a patient, using the median as cut-off to determine high and low groups.

1027

Methods and Extended Data References

27.   Raza, S. E. A. et al. Micro-Net: A unified model for segmentation of various objects in microscopy images. Med. Image Anal. 52, 160–173 (2019).

28.   Sirinukunwattana, K. et al. Locality Sensitive Deep Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology Images. IEEE Trans. Med. Imaging 35, 1196–1206 (2016).

29.   Abadi, M. et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. (2016).

30.   Danaher, P. et al. Gene expression markers of Tumor Infiltrating Leukocytes. J. Immunother. Cancer 5, 18 (2017).

31.   HOLZINGER, A. et al. Monoclonal Antibody to Thyroid Transcription Factor-1: Production, Characterization, and Usefulness in Tumor Diagnosis. Hybridoma 15, 49–53 (1996).

32.   Matoso, A. et al. Comparison of thyroid transcription factor-1 expression by 2 monoclonal antibodies in pulmonary and nonpulmonary primary tumors. Appl. Immunohistochem. Mol. Morphol. AIMM 18, 142–9 (2010).

33.   Pelosi, G. et al. ΔNp63 (p40) and Thyroid Transcription Factor-1 Immunoreactivity on Small Biopsies or Cellblocks for Typing Non-small Cell Lung Cancer: A Novel Two-Hit, Sparing-Material Approach. J. Thorac. Oncol. 7, 281–290 (2012).

34.   Heindl, A. et al. Relevance of Spatial Heterogeneity of Immune Infiltration for Predicting Risk of Recurrence After Endocrine Therapy of ER+ Breast Cancer. JNCI J. Natl. Cancer Inst. 110, (2018).

35.   Heindl, A. et al. Microenvironmental niche divergence shapes BRCA1-dysregulated ovarian cancer morphological plasticity. Nat. Commun. 9, 3917 (2018).

36.   Nawaz, S., Heindl, A., Koelble, K. & Yuan, Y. Beyond immune density: critical role of spatial heterogeneity in estrogen receptor-negative breast cancer. Mod. Pathol. 28, 766–777 (2015).

37.   Dubuc, Quiniou, Roques-Carmes, Tricot & Zucker. Evaluating the fractal dimension of profiles. Phys. Rev. A, Gen. Phys. 39, 1500–1512 (1989).

38.   MOISY, F. & JIMÉNEZ, J. Geometry and clustering of intense structures in isotropic turbulence. J. Fluid Mech. 513, 111–133 (2004).

39.   Francis, K. & Palsson, B. O. Effective intercellular communication distances are determined by the relative time constants for cyto/chemokine secretion and diffusion. Proc. Natl. Acad. Sci. U. S. A. 94, 12258–62 (1997).

1044

1045

References

1. Rosenthal, R. *et al.* Neoantigen-directed immune escape in lung cancer evolution. *Nature* 1 (2019). doi:10.1038/s41586-019-1032-7

2. Morris, L. G. T. & Chan, T. A. Lung Cancer Evolution: What's Immunity Got to Do with It? *Cancer Cell* **35**, 711–713 (2019).

3. Morris, L. G. T. *et al.* Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. *Oncotarget* **7**, 10051–10063 (2016).

4. Milo, I. *et al.* The immune system profoundly restricts intratumor genetic heterogeneity. *Sci. Immunol.* **3**, (2018).

5. Jia, Q. *et al.* Local mutational diversity drives intratumoral immune heterogeneity in non-small cell lung cancer. *Nat. Commun.* **9**, (2018).

6. Jamal-Hanjani, M. *et al.* Tracking the Evolution of Non–Small-Cell Lung Cancer. *N. Engl. J. Med.* **376**, 2109–2121 (2017).

7. Hendry, S. *et al.* Assessing Tumor-Infiltrating Lymphocytes in Solid Tumors. *Adv. Anat. Pathol.* **24**, 311–335 (2017).

8. Moore, D. A. *et al.* In situ growth in early lung adenocarcinoma may represent precursor growth or invasive clone outgrowth—a clinically relevant distinction. *Mod. Pathol.* 1 (2019). doi:10.1038/s41379-019-0257-1

9. Whittaker, K. A. & Rynearson, T. A. Evidence for environmental and ecological selection in a microbe with no geographic limits to gene flow. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 2651–2656 (2017).

10. Shafer, A. B. A. & Wolf, J. B. W. Widespread evidence for incipient ecological speciation: A meta-analysis of isolation-by-ecology. *Ecology Letters* **16**, 940–950 (2013).

11. Costa, A. *et al.* Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. *Cancer Cell* **33**, 463-479.e10 (2018).

12. Öhlund, D. *et al.* Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J. Exp. Med.* **214**, 579–596 (2017).

13. Salmon, H. *et al.* Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. *J. Clin. Invest.* **122**, 899–910 (2012).

14. Thomas, D. A. & Massagué, J. TGF-β directly targets cytotoxic T cell functions during tumor evasion of immune surveillance. *Cancer Cell* **8**, 369–380 (2005).

15. Joyce, J. A. & Fearon, D. T. T cell exclusion, immune privilege, and the tumor microenvironment. *Science (80-.).* **348**, 74–80 (2015).

16. Sorokin, L. The impact of the extracellular matrix on inflammation. *Nat. Rev. Immunol.* **10**, 712–723 (2010).

17. Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer–immune set point. *Nature* **541**, 321–330 (2017).

1085   18.   McGranahan, N. *et al.* Allele-Specific HLA Loss and Immune Escape in Lung Cancer
1086   Evolution. *Cell* **171**, 1259-1271.e11 (2017).
1087   19.   McGranahan, N. *et al.* Clonal neoantigens elicit T cell immunoreactivity and
1088   sensitivity to immune checkpoint blockade. *Science (80-.).* **351**, (2016).
1089   20.   Yuan, Y. Modelling the spatial heterogeneity and molecular correlates of
1090   lymphocytic infiltration in triple-negative breast cancer. *J. R. Soc. Interface* **12**,
1091   20141153 (2015).
1092   21.   Thomas, A., Liu, S. V., Subramaniam, D. S. & Giaccone, G. Refining the treatment of
1093   NSCLC according to histological and molecular subtypes. *Nat. Rev. Clin. Oncol.* **12**,
1094   511–526 (2015).
1095   22.   Hammerman, P. S. *et al.* Comprehensive genomic characterization of squamous cell
1096   lung cancers. *Nature* **489**, 519–525 (2012).
1097   23.   Collisson, E. A. *et al.* Comprehensive molecular profiling of lung adenocarcinoma:
1098   The cancer genome atlas research network. *Nature* **511**, 543–550 (2014).
1099   24.   Keren, L. *et al.* A Structured Tumor-Immune Microenvironment in Triple Negative
1100   Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell* **174**, 1373-1387.e19
1101   (2018).
1102   25.   Giesen, C. *et al.* Highly multiplexed imaging of tumor tissues with subcellular
1103   resolution by mass cytometry. *Nat. Methods* **11**, 417–422 (2014).
1104   26.   Goltsev, Y. *et al.* Deep Profiling of Mouse Splenic Architecture with CODEX
1105   Multiplexed Imaging. *Cell* **174**, 968-981.e15 (2018).
1106   27.   Raza, S. E. A. *et al.* Micro-Net: A unified model for segmentation of various objects
1107   in microscopy images. *Med. Image Anal.* **52**, 160–173 (2019).
1108   28.   Sirinukunwattana, K. *et al.* Locality Sensitive Deep Learning for Detection and
1109   Classification of Nuclei in Routine Colon Cancer Histology Images. *IEEE Trans. Med.*
1110   *Imaging* **35**, 1196–1206 (2016).
1111   29.   Abadi, M. *et al.* TensorFlow: Large-Scale Machine Learning on Heterogeneous
1112   Distributed Systems. (2016).
1113   30.   Danaher, P. *et al.* Gene expression markers of Tumor Infiltrating Leukocytes. *J.*
1114   *Immunother. Cancer* **5**, 18 (2017).
1115   31.   HOLZINGER, A. *et al.* Monoclonal Antibody to Thyroid Transcription Factor-1:
1116   Production, Characterization, and Usefulness in Tumor Diagnosis. *Hybridoma* **15**,
1117   49– 53 (1996).
1118   32.   Matoso, A. *et al.* Comparison of thyroid transcription factor-1 expression by 2
1119   monoclonal antibodies in pulmonary and nonpulmonary primary tumors. *Appl.*
1120   *Immunohistochem. Mol. Morphol. AIMM* **18**, 142–9 (2010).
1121   33.   Pelosi, G. *et al.* ΔNp63 (p40) and Thyroid Transcription Factor-1 Immunoreactivity
1122   on Small Biopsies or Cellblocks for Typing Non-small Cell Lung Cancer: A Novel Two-
1123   Hit, Sparing-Material Approach. *J. Thorac. Oncol.* **7**, 281–290 (2012).

1124   34.   Heindl, A. *et al.* Relevance of Spatial Heterogeneity of Immune Infiltration for
1125   Predicting Risk of Recurrence After Endocrine Therapy of ER+ Breast Cancer. *JNCI J.*
1126   *Natl. Cancer Inst.* **110**, (2018).
1127   35.   Heindl, A. *et al.* Microenvironmental niche divergence shapes BRCA1-dysregulated
1128   ovarian cancer morphological plasticity. *Nat. Commun.* **9**, 3917 (2018).
1129   36.   Nawaz, S., Heindl, A., Koelble, K. & Yuan, Y. Beyond immune density: critical role of
1130   spatial heterogeneity in estrogen receptor-negative breast cancer. *Mod. Pathol.* **28**,
1131   766–777 (2015).
1132   37.   Dubuc, Quiniou, Roques-Carmes, Tricot & Zucker. Evaluating the fractal dimension
1133   of profiles. *Phys. Rev. A, Gen. Phys.* **39**, 1500–1512 (1989).
1134   38.   MOISY, F. & JIMÉNEZ, J. Geometry and clustering of intense structures in isotropic
1135   turbulence. *J. Fluid Mech.* **513**, 111–133 (2004).
1136   39.   Francis, K. & Palsson, B. O. Effective intercellular communication distances are
1137   determined by the relative time constants for cyto/chemokine secretion and
1138   diffusion. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12258–62 (1997).

Geospatial immune variability illuminates differential evolution of lung adenocarcinoma

TRACERx consortium member names

Charles Swanton (3,4,5), Mariam Jamal-Hanjani (3,5), John Le Quesne (10,11,15), Allan Hackshaw (12), Sergio A Quezada (13), Nicholas McGranahan (3,14), Rachel Rosenthal (3,4), Crispin T Hiley (3,4), Selvaraju Veeriah (3,4), David A Moore (3,6), Maise Al Bakir (4), Teresa Marafioti (6), Roberto Salgado (8,9), Yenting Ngai (12), Abigail Sharp (12), Cristina Rodrigues (12), Oliver Pressey (12), Sean Smith (12), Nicole Gower (12), Harjot Dhanda (12), Joan Riley (16), Lindsay Primrose (16), Luke Martinson (16), Nicolas Carey (16), Jacqui A Shaw (16), Dean Fennell (16,28), Gareth A Wilson (17), Nicolai J Birkbak (17), Thomas B K Watkins (17), Mickael Escudero (17), Aengus Stewart (17), Andrew Rowan (17), Jacki Goldman (17), Peter Van Loo (17), Richard Kevin Stone (17), Tamara Denner (17), Emma Nye (17), Sophia Ward (17), Emilia L Lim (17), Stefan Boeing (17), Maria Greco (17), Kevin Litchfield (17), Jerome Nicod (17), Clare Puttick (17), Katey Enfield (17), Emma Colliver (17), Brittany Campbell (17), Christopher Abbosh (18), Yin Wu (18), Marcin Skrzypski (18), Robert E Hynds (18), Andrew Georgiou (18), Mariana Werner Sunderland (18), James L Reading (18), Karl S Peggs (18), John A Hartley (18), Pat Gorman (18), Helen L Lowe (18), Leah Ensell (18), Victoria Spanswick (18), Angeliki Karamani (18), Dhruva Biswas (18), Maryam Razaq (18), Stephan Beck (18), Ariana Huebner (18), Michelle Dietzen (18), Cristina Naceur-Lombardelli (18), Mita Afroza Akther (18), Haoran Zhai (18), Nnennaya Kannu (18), Elizabeth Manzano (18), Supreet Kaur Bola (18), Ehsan Ghorani (18), Marc Robert de Massy (18), Elena Hoxha (18), Emine Hatipoglu (18), Stephanie Ogwuru (18), Benny Chain (18), Gillian Price (19), Sylvie DuboisMarshall (19), Keith Kerr (19), Shirley Palmer (19), Heather Cheyne (19), Joy Miller (19), Keith Buchan (19), Mahendran Chetty (19), Mohammed Khalil (19), Veni Ezhil (20), Vineet Prakash (20), Girija Anand (21), Sajid Khan (21), Kelvin Lau (22), Michael Sheaff (22), Peter Schmid (22), Louise Lim (22), John Conibear (22), Roland Schwarz (23,24,25), Jonathan Tugwood (26), Jackie Pierce (26), Caroline Dive (26,27), Ged Brady (26,27), Dominic G Rothwell (26,27), Francesca Chemi (26,27), Elaine Kilgour (26,27), Fiona Blackhall (27,30), Lynsey Priest (27,30), Matthew G Krebs (27,30), Philip Crosbie (27,51,52), Apostolos Nakas (28), Sridhar Rathinam (28), Louise Nelson (28), Kim Ryanna (28), Mohamad Tuffail (28), Amrita Bajaj (28), Jan Brozik (28), Fiona Morgan (29), Malgorzata Kornaszewska (29), Richard Attanoos (29), Haydn Adams (29), Helen Davies (29), Mathew Carter (30), Lindsay CR (30), Fabio Gomes (30), Zoltan Szallasi (31), Istvan Csabai (32), Miklos Diossy (32), Hugo Aerts (33,34), Alan Kirk (35), Mo Asif (35), John Butler (35), Rocco Bilanca (35), Nikos Kostoulas (35), Mairead MacKenzie (36), Maggie Wilcox (36), Sara

Busacca (37), Alan Dawson (37), Mark R Lovett (37), Michael Shackcloth (38), Sarah Feeney (38), Julius Asante-Siaw (38), John Gosney (39), Angela Leek (40), Nicola Totten (40), Jack Davies Hodgkinson (40), Rachael Waddington (40), Jane Rogan (40), Katrina Moore (40), William Monteiro (41), Hilary Marshall (41), Kevin G Blyth (42), Craig Dick (42), Andrew Kidd (42), Eric Lim (43), Paulo De Sousa (43), Simon Jordan (43), Alexandra Rice (43), Hilgardt Raubenheimer (43), Harshil Bhayani (43), Morag Hamilton (43), Lyn Ambrose (43), Anand Devaraj (43), Hema Chavan (43), Sofina Begum (43), Aleksander Mani (43), Daniel Kaniu (43), Mpho Malima (43), Sarah Booth (43), Andrew G Nicholson (43), Nadia Fernandes (43), Jessica E Wallen (43), Pratibha Shah (43), Sarah Danson (44), Jonathan Bury (44), John Edwards (44), Jennifer Hill (44), Sue Matthews (44), Yota Kitsanta (44), Jagan Rao (44), Sara Tenconi (44), Laura Socci (44), Kim Suvarna (44), Faith Kibutu (44), Patricia Fisher (44), Robin Young (44), Joann Barker (44),

Fiona Taylor (44), Kirsty Lloyd (44), Teresa Light (45), Tracey Horey (45), Dionysis PapadatosPastos (45, 47), Peter Russell (45), Sara Lock (46), Kayleigh Gilbert (46), David Lawrence (47), Martin Hayward (47), Nikolaos Panagiotopoulos (47), Robert George (47), Davide Patrini

(47), Mary Falzon (47), Elaine Borg (47), Reena Khiroya (47), Asia Ahmed (47), Magali Taylor (47), Junaid Choudhary (47), Penny Shaw (47), Sam M Janes (47), Martin Forster (47), Tanya Ahmad (47), Siow Ming Lee (47), Javier Herrero (47), Dawn Carnell (47), Ruheena Mendes (47), Jeremy George (47), Neal Navani (47), Marco Scarci (47), Elisa Bertoja (47), Robert CM Stephens (47), Emilie Martinoni Hoogenboom (47), James W Holding (47), Steve Bandula (47), Babu Naidu (48), Gerald Langman (48), Andrew Robinson (48), Hollie Bancroft (48), Amy Kerr (48), Salma Kadiri (48), Charlotte Ferris (48), Gary Middleton (48), Madava Djearaman (48), Akshay Patel (48), Christian Ottensmeier (49), Serena Chee (49), Benjamin Johnson (49), Aiman Alzetani (49), Emily Shaw (49), Jason Lester (50), Yvonne Summers (51), Raffaele Califano (51), Paul Taylor (51), Rajesh Shah (51), Piotr Krysiak (51), Kendadai Rammohan (51), Eustace Fontaine (51), Richard Booton (51), Matthew Evison (51), Stuart Moss (51), Juliette Novasio (51), Leena Joseph (51), Paul Bishop (51), Anshuman Chaturvedi (51), Helen Doran (51), Felice Granato (51), Vijay Joshi (51), Elaine Smith (51), Angeles Montero (51)

TRACERx consortium affiliations

(3)   Cancer Research UK Lung Cancer Centre of Excellence, University College London Cancer Institute, London, UK

(4)   Cancer Evolution and Genome Instability Laboratory, The Francis Crick Institute, London, UK

(5)   Department of Medical Oncology, University College London Hospitals NHS Foundation Trust, London, UK

1217	(6)	Department of Cellular Pathology, University College London, University College Hospital, London, UK
1218		
1219	(7)	Translational Immune Oncology Group, Centre for Molecular Medicine, Royal Marsden Hospital NHS Trust, London, UK
1220		
1221	(8)	Department of Pathology, GZA-ZNA-Ziekenhuizen, Antwerp, Belgium
1222	(9)	Division of Research, Peter MacCallum Cancer Centre, University of Melbourne, Melbourne, Victoria, Australia
1223		
1224	(10)	MRC Toxicology Unit, Lancaster Road, University of Cambridge, Leicester, UK
1225	(11)	Leicester Cancer Research Centre, University of Leicester, Leicester, UK
1226	(12)	Cancer Research UK & University College London Cancer Trials Centre, University College London, London, UK
1227		
1228	(13)	Cancer Immunology Unit, University College London Cancer Institute, London, UK
1229	(14)	Cancer Genome Evolution Research Group, University College London Cancer Institute, University College London, London, UK
1230		
1231	(15)	Glenfield Hospital, University Hospitals Leicester NHS Trust, Groby Road, Leicester, UK
1232	(16)	Cancer Research Centre, University of Leicester, Leicester, United Kingdom
1233	(17)	The Francis Crick Institute, London, United Kingdom
1234	(18)	University College London Cancer Institute, London, United Kingdom
1235	(19)	Aberdeen Royal Infirmary, Aberdeen, United Kingdom
1236	(20)	Ashford and St Peter's Hospitals NHS Foundation Trust
1237	(21)	Barnet & Chase Farm Hospitals, United Kingdom
1238	(22)	Barts Health NHS Trust
1239	(23)	Berlin Institute for Medical Systems Biology, Max Delbrueck Center for Molecular Medicine, Berlin, Germany
1240		
1241	(24)	German Cancer Consortium (DKTK), partner site Berlin
1242	(25)	German Cancer Research Center (DKFZ), Heidelberg
1243	(26)	Cancer Research UK Manchester Institute, University of Manchester, Manchester, UK
1244	(27)	Cancer Research UK Lung Cancer Centre of Excellence, University of Manchester, Manchester, UK
1245		
1246	(28)	Leicester University Hospitals, Leicester, United Kingdom
1247	(29)	Cardiff & Vale University Health Board, Cardiff, Wales
1248	(30)	Christie NHS Foundation Trust, Manchester, United Kingdom

1249   (31)   Danish Cancer Society Research Center, Copenhagen, Denmark

1250   (32)   Department of Physics of Complex Systems, ELTE Eötvös Loránd University, Budapest,
1251   Hungary

1252   (33)   Departments of Radiation Oncology and Radiology, Dana Farber Cancer Institute, Brigham
1253   and Women's Hospital, Harvard Medical School, Boston, MA, USA

1254   (34)   Department of Radiology, Netherlands Cancer Institute, Amsterdam, The Netherlands

1255   (35)   Golden Jubilee National Hospital

1256   (36)   Independent Cancer Patients Voice

1257   (37)   University of Leicester, Leicester, United Kingdom

1258   (38)   Liverpool Heart and Chest Hospital NHS Foundation Trust

1259   (39)   Royal Liverpool University Hospital

1260   (40)   Manchester Cancer Research Centre Biobank, Manchester, United Kingdom

1261   (41)   National Institute for Health Research Leicester Respiratory Biomedical Research Unit,
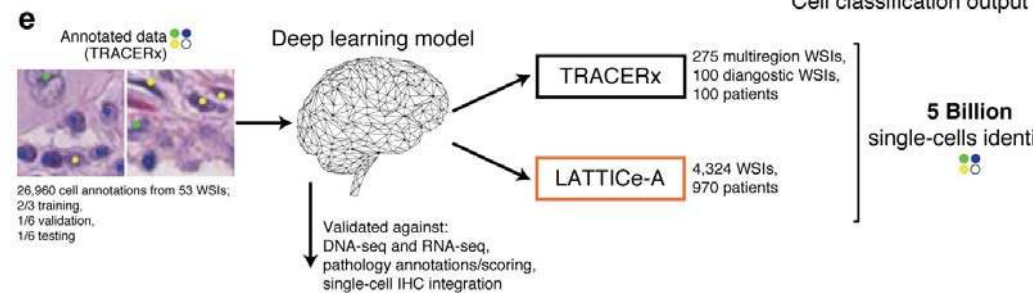1262   Leicester, United Kingdom

1263   (42)   NHS Greater Glasgow and Clyde

1264   (43)   Royal Brompton and Harefield NHS Foundation Trust

1265   (44)   Sheffield Teaching Hospitals NHS Foundation Trust

1266   (45)   The Princess Alexandra Hospital NHS Trust

1267   (46)   The Whittington Hospital NHS Trust, London, United Kingdom

1268   (47)   University College London Hospitals, London, United Kingdom

1269   (48)   University Hospital Birmingham NHS Foundation Trust, Birmingham, United Kingdom

1270   (49)   University Hospital Southampton NHS Foundation Trust

1271   (50)   Velindre Cancer Centre, Cardiff, Wales

1272   (51)   Wythenshawe Hospital, Manchester University NHS Foundation Trust

1273   (52)   Division of Infection, Immunity and Respiratory Medicine, University of Manchester,
1274   Manchester, UK

1275

TRACERx multiregion sampling

Diagnostic block

R1 R3 R4 R2

FFPE/H&E

A/RNA

Diagnostic

IHC
n = 100

H&E
n = 100

Regional

H&E

H&E

H&E
n = 275

b Tissue segmentation

Convolution & Max-pooling   Convolution   Convolution, Max-pooling & Deconvolution

subsample & concatenate   intermediate connections

Cell detection and classification

Convolution   Pooling Convolution Pooling Convolution   Fully connected   SC1   SC2

Cell classification output

e Annotated data (TRACERx)

26,960 cell annotations from 53 WSIs;
2/3 training,
1/6 validation,
1/6 testing

Deep learning model

Validated against:
DNA-seq and RNA-seq,
pathology annotations/scoring,
single-cell IHC integration

TRACERx   275 multiregion WSIs,
100 diagnostic WSIs,
100 patients

LATTICe-A   4,324 WSIs,
970 patients

5 Billion
single-cells identi

H&E

Cancer

Stromal

Lymphocyte

Other

IHC

CD8+

FOXP3-
FOXP3+

Other

CD4+FOXP3-

CD4+FOXP3+

f

IHC   H&E-IHC overlay   H&E cell classification

TTF1+

• TTF1+ expression

CD45+

• CD45+ expression   • Cancer • Stromal • Lymphocyte ○ Other

g rho = 0.93, P < 2.2e-16

TTF1+ cell / 100μm²

0.8
0.6
0.4
0.2
0.0

0.0   0.2   0.4   0.6

H&E-based cancer cell / 100μm

h rho = 0.83, P < 2.2e-16

CD45+ cell / 100μm²

0.6
0.4
0.2
0.0

0.0   0.1   0.2   0.3

H&E-based lymphocyte cell / 100μ

Immune cold

Immune hot

=275 region; 85 patients

LUAD   LUSC   Oth

50

40

30

20

10

CRUK0034 CRUK0037 CRUK0022 CRUK0007 CRUK0025 CRUK0023 CRUK0050 CRUK0051 CRUK0003 CRUK0041 CRUK0036 CRUK0052 CRUK0048 CRUK0029 CRUK0010 CRUK0042 CRUK0006 CRUK0015 CRUK0001 CRUK0018 CRUK0061 CRUK0035 CRUK0031 CRUK0002 CRUK0044 CRUK0047 CRUK0046 CRUK0055 CRUK0009 CRUK0038 CRUK0024 CRUK0013 CRUK0004 CRUK0053 CRUK0032 CRUK0016 CRUK0021 CRUK0028 CRUK0008 CRUK0020 CRUK0039 CRUK0017 CRUK0012 CRUK0057 CRUK0045 CRUK0026 CRUK0027 CRUK0030 CRUK0087 CRUK0071 CRUK0063 CRUK0077 CRUK0089 CRUK0070 CRUK0092 CRUK0062 CRUK0067 CRUK0065 CRUK0086 CRUK0076 CRUK0084 CRUK0079 CRUK0081 CRUK0083 CRUK0075 CRUK0085 CRUK0074 CRUK0091 CRUK0090 CRUK0066 CRUK0073 CRUK0078 CRUK0082 CRUK0069 CRUK0088 CRUK0072 CRUK0068 CRUK0100 CRUK0094 CRUK0086 CRUK0099

08+%

0P3 %

P3+%

Immune signature using Danaher et al

Immune cold

Immune intermediate

Immune hot

Lymphocyte% CD8 CD4 B cell CD45 Cyto Dend CD8 exhausted Macrophage Neutrophil NK CD56⁻ NK T cells T_H1 T_reg

Immune phenotype
- Hot
- Intermediate
- Cold

CD8+%
20
15
10
5
0

CD4+FOXP3 %
30
20
10
0

CD4+FO
20
15
10
0

d

Estimated infiltration

6

4

2

0

*** *** *** *** *** * *** ** ** *** *** *** *** *

CD8 CD4 B cell CD45 Cyto Dendritic CD8 exhausted Macrophage Neutrophil NK CD56⁻ NK T cells T_H1 T_reg

Immune signature using Danaher et al

Corrected
*** p < 0.0001
** p < 0.01
* p < 0.05

Immune phenotype
- cold
- hot

b

Distance between dominant clones to the last common ancestor of regions pair

0.00036

cold-cold   hot-hot
Immune phenotype

0.019

cold-cold   hot-hot
Immune phenotype

c

CRUK0036, **immune cold pair**

R1   R3

CRUK0024, **immune hot pair**

R3   R6

● Dominant clone (CCF ≥ 75%)

● Minor clone (CCF < 75%)

● Undetected

● Most recent common ancestor of all cancer cells

↗ ↖ Last common ancestor of region pair

d TRACERx LUAD and LUSC
n = 79, WSIs/regions = 249

Disease-free Survival

logrank p = 0.0035

0   200 400 600 800 1000 1200 1400
Days to Death or Recurrence

Number at Risk

High	29	23	17	11	10	8	4	0
Low	50	46	44	41	38	27	11	0

e LATTICe-A LUAD
n = 970, WSIs/samples = 4,324

Disease-free Survival

logrank p = 1.95x10^{-5}

0   1000 2000 3000 4000 5000 6000
Days to Death or Recurrence

Number at Risk

High	427	121	39	24	11	4	0
Low	543	247	103	56	29	13	1

— <=1 cold region (median)
— >1 cold region

f TRACERx

N cold regions	(N=50)			
	>1 (N=29)	2.65 (1.184 – 5.95)		0.018 *
Age	(N=79)	1.03 (0.978 – 1.09)		0.262
Sex	Female (N=27)			
	Male (N=52)	0.47 (0.206 – 1.08)		0.074
Pack years	(N=79)	1.00 (0.981 – 1.01)		0.69
Histology	LUAD (N=49)			
	LUSC (N=30)	1.30 (0.530 – 3.17)		0.569
Stage	IA (N=16)			
	IB (N=30)	1.21 (0.350 – 4.21)		0.761
	IIA (N=12)	3.59 (0.731 – 17.63)		0.116
	IIB (N=9)	6.41 (1.150 – 35.69)		0.034 *
	III (N=12)	6.52 (1.585 – 26.83)		0.009 **
Adjuvant therapy	No (N=54)			
	Yes (N=25)	0.40 (0.151 – 1.04)		0.059
Clonal Neo.	Low (N=58)			
	High (N=21)	0.24 (0.076 – 0.77)		0.016 *
N total regions	(N=79)	1.22 (0.858 – 1.73)		0.271

Events: 34; Global p-value (Log-Rank): 0.00014034
AIC: 245.75; Concordance Index: 0.82

0.1 0.2 0.5 1 2 5 10 20 50
Hazard ratio

g LATTICe-A

N cold samples	<=1 (N=372)			
	>1 (N=279)	1.48 (1.17 – 1.9)		<0.001 ***
Age	(N=651)	0.99 (0.98 – 1.0)		0.198
Sex	Female (N=377)			
	Male (N=274)	1.21 (0.96 – 1.5)		0.105
Pack years	(N=651)	1.00 (1.00 – 1.0)		0.08
Stage	IA (N=165)			
	IB (N=161)	1.60 (1.09 – 2.3)		0.016 *
	IIA (N=101)	3.30 (2.23 – 4.9)		<0.001 ***
	IIB (N=82)	3.76 (2.49 – 5.7)		<0.001 ***
	III (N=142)	4.59 (3.12 – 6.7)		<0.001 ***
Adjuvant therapy	No (N=533)			
	Yes (N=118)	0.85 (0.64 – 1.1)		0.291
N total samples	(N=651)	1.00 (0.94 – 1.1)		0.979

Events: 328; Global p-value (Log-Rank): 2.0234e-20
AIC: 3727.49; Concordance Index: 0.68

1   2   5
Hazard ratio

a

b

Fractal dimension

0.00036

cold   hot
Immune phenotype

c

Fractal dimension

0.0057

No loss   Loss
LOH: any HLA type

H&E

H&E-based spatial modelling of lymphocytes

IHC-derived immune subsets

e

Adjacent–tumor lymphocytes/stroma

0.0074

>=Median   <Median
Clonal neoantigens

● Intra-tumor lymphocyte
● Adjacent-tumor lymphocyte
● Distal-tumor lymphocyte

● CD8+
● CD4+FOXP3-
● CD4+FOXP3+

f

CD8,CD4,FOXP3

Spatial alignment

H&E

Composite

Project IHC-derived cells onto H&E

+ CD8+
+ CD4+FOXP3-   ] IHC
+ CD4+FOXP3+

Cancer density/contour plot from H&E

g

$CD8^+/CD4^+FOXP3^+$

0.3

0.018

0.0095

Adjacent   Distal   Intra
Spatial compartment

a TRACERx histology CONSORT diagram

TRACERx 100 patient cohort [2017, Jamal-Hanjani]
n=100 patients; 303 regions

Multi-region sampling

Excluded 15 patients (insufficient amount of tissue)

Ⓐ Ⓒ H&E
n=85 patients; 275 regions

Common exome samples: n= 83 patients; 238 regions
Common RNA samples: n=56 patients; 142 regions
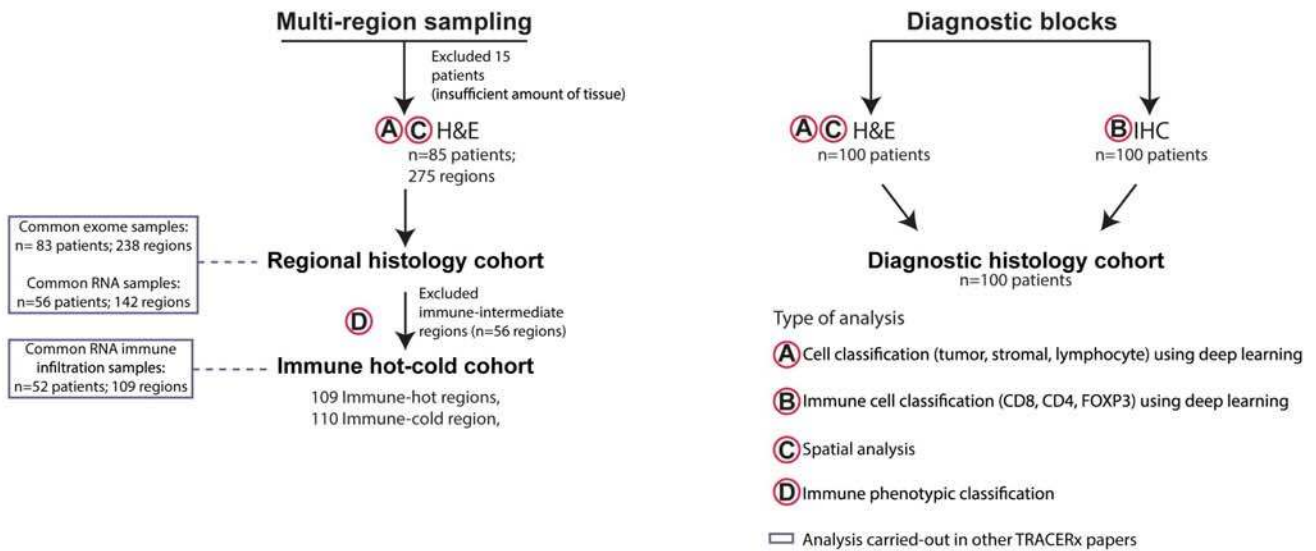
Regional histology cohort

Excluded immune-intermediate regions (n=56 regions)

Common RNA immune infiltration samples: n=52 patients; 109 regions

Ⓓ **Immune hot-cold cohort**

109 Immune-hot regions,
110 Immune-cold region,

Diagnostic blocks

Ⓐ Ⓒ H&E
n=100 patients

Ⓑ IHC
n=100 patients

Diagnostic histology cohort
n=100 patients

Type of analysis

Ⓐ Cell classification (tumor, stromal, lymphocyte) using deep learning

Ⓑ Immune cell classification (CD8, CD4, FOXP3) using deep learning

Ⓒ Spatial analysis

Ⓓ Immune phenotypic classification

☐ Analysis carried-out in other TRACERx papers

b TRACERx patient characteristics

LUAD   LUSC   Other

Histology
Stage
Pack years
Recurr./death
N regions
N RNA regions
N hist. regions
N cold regions

Stage
- IA
- IB
- IIA
- IIB
- IIIA
- IIIB

Pack years
150
100
50
0

Recurrence/death
- Yes
- No

Number of total/RNA/histology regions
1 2 3 4 5 6 7 8

Number of immune-cold regions
0 1 2 3 4 5

NA

c Validation CONSORT diagram

LATTICe-A cohort [2019, Moore]
n=994 patients; 5,001 samples

Multi-sample diagnostic blocks

Excluded:
664 samples (lymph node tissue),
24 patients (no clinical data)

Ⓐ Ⓒ H&E
n=970 patients;
4,324 samples

Multi-sample histology cohort

Excluded immune-intermediate samples (n=876 regions)

Ⓓ **Immune hot-cold cohort**

1,716 Immune-hot samples,
1,732 Immune-cold samples

d Demographic and clinical overview

TRACERx cohort
Sex Female Male

LATTICe-A cohort

Pathological stage
- IA
- IB
- IIA
- IIB
- III

a TRACERx, diagnostic samples (n=98)

rho = 0.31, *P* = 0.017
rho = 0.38, *P* = 0.032
rho = 0.36, *P* = 0.00022

Pathology TIL estimates

b TRACERx, regional histology and CNA samples (n=238)

rho = 0.32, *P* = 0.00029
rho = 0.35, *P* = 0.00057
rho = 0.36, *P* = 7.3e-09

Tumor%

ASCAT Purity

rho = 0.32, *P* = 0.0
rho = 0.35, *P* = 0.0
rho = 0.36, *P* = 7.7

Tumor%

VAF Purity

c TRACERx, regional histology and RNA-seq samples (n=142)

rho = 0.25, *P* = 0.029
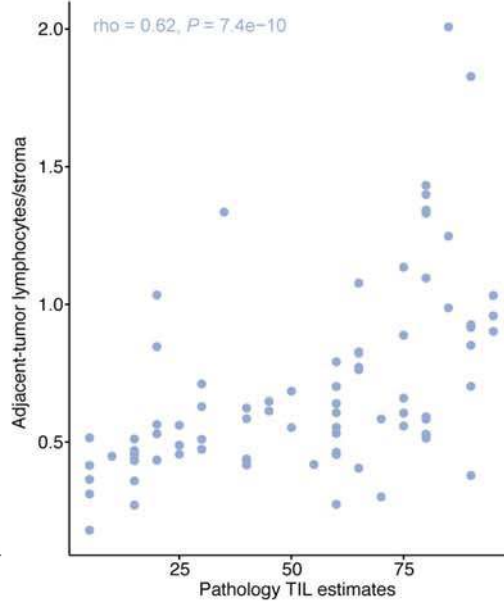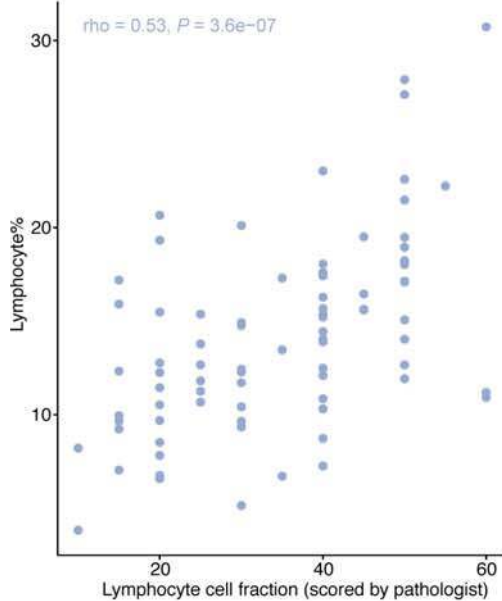rho = 0.43, *P* = 0.0023
rho = 0.44, *P* = 4e-08

CD8 score (Danaher et. al. signature)

d TRACERx, diagnostic samples (n=100)

rho = 0.56, *P* = 2.3e-06
rho = 0.43, *P* = 0.013
rho = 0.53, *P* = 1.8e-08

IHC CD8%

H&E Lymphocyte%

f

IHC

H&E
Serial staining

e LATTICe-A, diagnostic samples (n=80)

rho = 0.53, *P* = 3.6e-07

Lymphocyte%

Lymphocyte cell fraction (scored by pathologist)

rho = 0.62, *P* = 7.4e-10

Adjacent-tumor lymphocytes/stroma

Pathology TIL estimates

g

rho=0.74, *P* <2.2e-16

SMA⁺ cell / 100μm²

H&E-based stromal cell / 100μm²

LUAD
LUSC
Other

b

Deep learning

	Hot	Cold
RNA-seq High	35	18
Low	13	43

p=7.8e−06

c

Fraction of immune hotspot

0.011

Discrepant   In agreement

Classification agreement

d

Lymphocyte%   ****   cold   hot

Intra−tumor lym/tumor   ****   cold   hot

Adjacent−tumor lym/stroma   ****   cold   hot

ASCAT.purity   ****   cold   hot

H&E-deep learning region specific immune classification

Lymphocyte%   ****   low   high

Intra−tumor lym/tumor   ****   low   high

Adjacent−tumor lym/stroma   ***   low   high

ASCAT.purity   ****   low   high

RNA-seq regions specific immune classification

4.6e−08

Cold   Hot

Immune classification

a

Genomic distance vs Immune phenotype (cold-cold, hot-hot), p = 0.49

b

PHBR2 (using RNA-seq regional immune classification)

N immune-low regions (RNA-seq) (N=64) — HR 1.5 (1.2–1.8), p = 0.002 **
Events: 29; Global p-value (Log-Rank): 0.0029913
AIC: 200.7; Concordance Index: 0.61

N immune-low regions (RNA-seq) <=1 (N=42); >1 (N=22) — HR 2.8 (1.3–5.9), p = 0.008 **
Events: 29; Global p-value (Log-Rank): 0.0081002
AIC: 202.5; Concordance Index: 0.59

N immune-high regions (RNA-seq) (N=64) — HR 1.1 (0.88–1.4), p = 0.336
Events: 29; Global p-value (Log-Rank): 0.3493
AIC: 206.84; Concordance Index: 0.56

c Multivariate analysis

TRACERx

N cold regions	(N=50)		
>1	(N=29)	2.90 (1.31–6.4)	0.009 **
Age	(N=79)	1.02 (0.97–1.1)	0.409
sex	Female (N=27)		
	Male (N=52)	0.47 (0.21–1.1)	0.071
Pack years	(N=79)	1.00 (0.98–1.0)	0.769
Histology	LUAD (N=49)		
	LUSC (N=30)	1.45 (0.62–3.4)	0.392
Stage	IA (N=16)		
	IB (N=30)	1.41 (0.42–4.8)	0.583
	IIA (N=12)	5.00 (1.14–22.0)	0.033 *
	IIB (N=9)	10.00 (2.19–45.7)	0.003 **
	III (N=12)	8.03 (2.04–31.6)	0.003 **
Adjuvant therapy	No (N=54)		
	Yes (N=25)	0.38 (0.14–1.0)	0.056
Clonal Neo.	Low (N=58)		
	High (N=21)	0.22 (0.07–0.7)	0.01 **

Events: 34; Global p-value (Log-Rank): 0.00011444
AIC: 244.97; Concordance Index: 0.81

LATTICe-A

N cold samples	<=1 (N=372)		
>1	(N=279)	1.48 (1.18–1.8)	<0.001 ***
Age	(N=651)	0.99 (0.98–1.0)	0.198
Sex	Female (N=377)		
	Male (N=274)	1.21 (0.96–1.5)	0.104
Pack years	(N=651)	1.00 (1.00–1.0)	0.08
Stage	IA (N=165)		
	IB (N=161)	1.60 (1.09–2.3)	0.016 *
	IIA (N=101)	3.30 (2.24–4.9)	<0.001 ***
	IIB (N=82)	3.77 (2.50–5.7)	<0.001 ***
	III (N=142)	4.60 (3.15–6.7)	<0.001 ***
Adjuvant therapy	No (N=533)		
	Yes (N=118)	0.85 (0.64–1.1)	0.291

Events: 328; Global p-value (Log-Rank): 5.4162e–21
AIC: 3725.49; Concordance Index: 0.68

d Multivariate analysis (including tumor size)

TRACERx

N cold regions	<=1 (N=50)		
>1	(N=29)	2.59 (1.138–5.89)	0.023 *
Age	(N=79)	1.03 (0.978–1.08)	0.268
Sex	Female (N=27)		
	Male (N=52)	0.46 (0.204–1.05)	0.066
Pack years	(N=79)	1.00 (0.984–1.01)	0.834
Histology	LUAD (N=49)		
	LUSC (N=30)	1.26 (0.525–3.03)	0.603
Stage	IA (N=16)		
	IB (N=30)	1.18 (0.330–4.23)	0.798
	IIA (N=12)	3.46 (0.673–17.75)	0.137
	IIB (N=9)	6.07 (1.023–35.99)	0.047 *
	III (N=12)	4.86 (0.926–25.54)	0.062
Adjuvant therapy	No (N=54)		
	Yes (N=25)	0.51 (0.171–1.51)	0.222
Clonal Neo.	Low (N=58)		
	High (N=21)	0.23 (0.074–0.74)	0.013 *
Tumor size	(N=79)	1.01 (0.990–1.03)	0.294

Events: 34; Global p-value (Log-Rank): 0.00014879
AIC: 245.91; Concordance Index: 0.82

LATTICe-A

N cold samples	<=1 (N=365)		
>1	(N=276)	1.46 (1.17–1.8)	0.001 **
Age	(N=641)	0.99 (0.98–1.0)	0.21
Sex	Female (N=373)		
	Male (N=268)	1.21 (0.96–1.5)	0.109
Pack years	(N=641)	1.00 (1.00–1.0)	0.06
Stage	IA (N=164)		
	IB (N=157)	1.50 (1.01–2.2)	0.042 *
	IIA (N=100)	2.96 (1.98–4.4)	<0.001 ***
	IIB (N=80)	2.95 (1.81–4.8)	<0.001 ***
	III (N=140)	4.23 (2.83–6.3)	<0.001 ***
Adjuvant therapy	No (N=523)		
	Yes (N=118)	0.83 (0.61–1.1)	0.205
Tumor size	(N=641)	1.00 (1.00–1.0)	0.048 *

Events: 324; Global p-value (Log-Rank): 1.5929e–21
AIC: 3664.75; Concordance Index: 0.68

e

TRACERx

Immune feature (Hazard ratio):
- N cold regions
- >1 cold regions
- N cold regions/total
- N hot regions
- Diagnostic lymphocyte%
- Regional lymphocyte% (min)
- Regional lymphocyte% (mean)
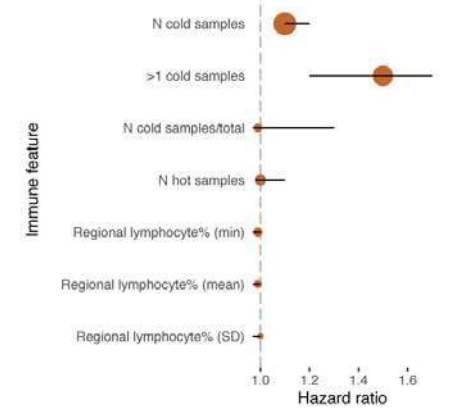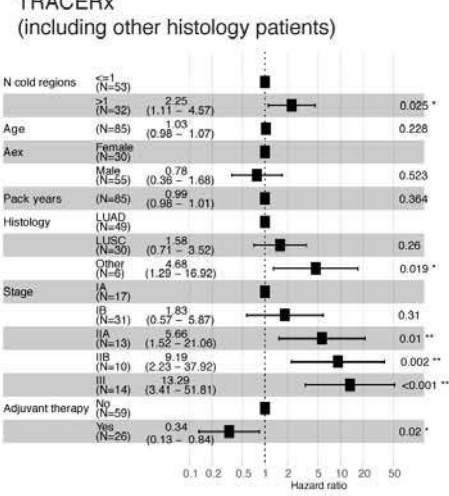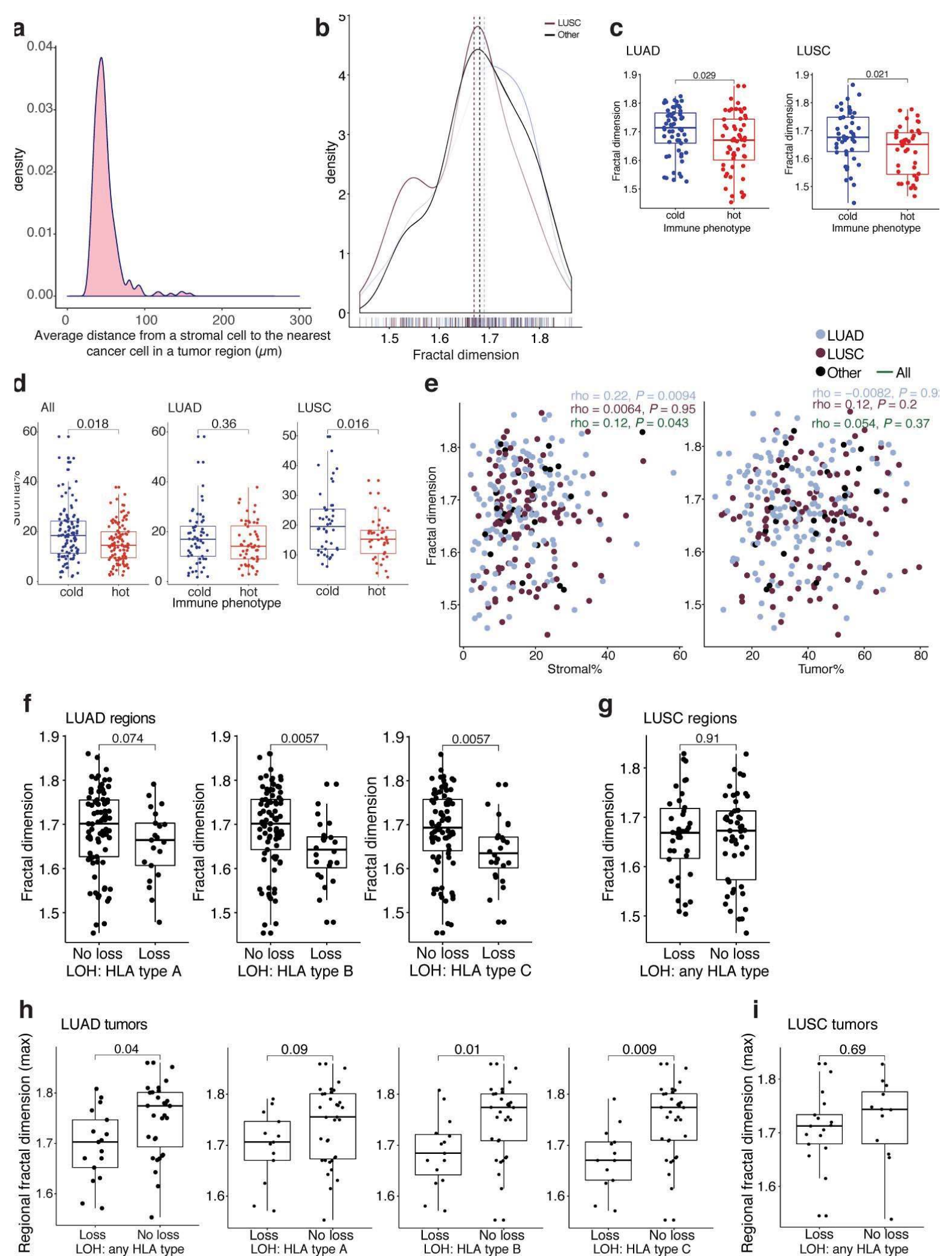- Regional lymphocyte% (SD)
- CD8⁺%
- CD4⁺FOXP3⁺%
- CD8⁻% / CD4⁺FOXP3⁺%
- CD4⁺FOXP3⁻% / CD4⁺FOXP3⁺%

LATTICe-A

Immune feature (Hazard ratio):
- N cold samples
- >1 cold samples
- N cold samples/total
- N hot samples
- Regional lymphocyte% (min)
- Regional lymphocyte% (mean)
- Regional lymphocyte% (SD)

f TRACERx

N cold	(N=79)	1.4 (1.10–1.7)	0.006 **
Subclonal CNA	(N=79)	1.00 (0.98–1.0)	0.975

N cold	(N=79)	1.68 (1.23–2.3)	<0.001 ***
ASCAT purity	(N=79)	0.59 (0.027–12.8)	0.715

N cold	(N=79)	1.47 (1.08–2.0)	0.015 *
Tumor%	(N=79)	0.99 (0.96–1.0)	0.634

g TRACERx (including other histology patients)

Disease-free Survival vs Days to Death or Recurrence
- <=1 cold region (median)
- >1 cold region

p = 0.025, HR 2.25 (1.11–4.57) adjusted for clinical parameters

Number at Risk
High	32	25	18	11	10	8	4	0
Low	53	47	45	42	39	28	12	0

h TRACERx (including other histology patients)

N cold regions	<=1 (N=53)		
>1	(N=32)	2.25 (1.11–4.57)	0.025 *
Age	(N=85)	1.03 (0.98–1.07)	0.228
Aex	Female (N=30)		
	Male (N=55)	0.78 (0.36–1.68)	0.523
Pack years	(N=85)	0.99 (0.98–1.01)	0.364
Histology	LUAD (N=49)		
	LUSC (N=30)	1.58 (0.71–3.52)	0.26
	Other (N=6)	4.68 (1.29–16.92)	0.019 *
Stage	IA (N=17)		
	IB (N=31)	1.83 (0.57–5.87)	0.31
	IIA (N=13)	5.66 (1.52–21.06)	0.01 **
	IIB (N=10)	9.19 (2.23–37.92)	0.002 **
	III (N=14)	13.29 (3.41–51.81)	<0.001 ***
Adjuvant therapy	No (N=59)		
	Yes (N=26)	0.34 (0.13–0.84)	0.02 *

H&E   IHC   H&E   IHC

c

d

Spatial TIL class