This is a repository copy of *Geospatial immune variability illuminates differential evolution of lung adenocarcinoma*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/212692/

Version: Accepted Version

Geospatial immune variability illuminates differential evolution of lung adenocarcinoma

Khalid AbdulJabbar[1,2,^], Shan E Ahmed Raza[1,2,^], Rachel Rosenthal[3,4,#], Mariam Jamal-Hanjani[3,5,#], Selvaraju Veeriah[3,4,#], Ayse Akarca[6], Tom Lund[7], David A. Moore[3,6], Roberto Salgado[8,9], Maise Al Bakir[4], Luis Zapata[1,2], Crispin T. Hiley[3,4], Leah Officer[10], Marco Sereno[11], Claire Rachel Smith[11], Sherene Loi[9], Allan Hackshaw[12], Teresa Marafioti[6], Sergio A. Quezada[13], Nicholas McGranahan[3,14], John Le Quesne[10,11,15,*], TRACERx consortium, Charles Swanton[3,4,5,*] and Yinyin Yuan[1,2,*]

[^]: Joint first authors
[#]: Joint second authors
[*]: Joint corresponding authors

[1]Centre for Evolution and Cancer, The Institute of Cancer Research, London, UK
[2]Division of Molecular Pathology, The Institute of Cancer Research, London, UK
[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
[6]Department of Cellular Pathology, University College London, University College Hospital, London, UK
[7]Translational Immune Oncology Group, Centre for Molecular Medicine, Royal Marsden Hospital NHS Trust, London, UK
[8]Department of Pathology, GZA-ZNA-Ziekenhuizen, Antwerp, Belgium
[9]Division of Research, Peter MacCallum Cancer Centre, University of Melbourne, Melbourne, Victoria, Australia
[10]MRC Toxicology Unit, Lancaster Road, University of Cambridge, Leicester, UK
[11]Leicester Cancer Research Centre, University of Leicester, Leicester, UK
[12]Cancer Research UK & University College London Cancer Trials Centre, University College London, London, UK
[13]Cancer Immunology Unit, University College London Cancer Institute, London, UK
[14]Cancer Genome Evolution Research Group, University College London Cancer Institute, University College London, London, UK
[15]Glenfield 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.

Main Text

Using an artificial intelligence framework, we developed a generalizable deep learning pipeline to spatially profile immune infiltration and discover tumor topological determinants of immunosuppression in digital pathology. Convolutional neural networks were tailored for the analysis of NSCLC morphology using diverse histology samples in the multi-region TRACERx 100 cohort[6] to avoid overfitting (Methods). This approach enabled the spatial mapping of cancer cells, lymphocytes, stromal cells (fibroblasts and endothelial cells), and an "other" cell class (macrophages, pneumocytes and non-identifiable cells) in hematoxylin & eosin (H&E)-stained images (275 tumor regions from 85 patients and 100 diagnostic slides from all patients, Fig. 1a-c, CONSORT diagram Extended Data Fig. 1a-b, Supplementary Table 1). T cell subsets were also identified in CD4/CD8/FOXP3 immunohistochemistry (IHC) images for all 100 diagnostic samples (Fig. 1d).

This pipeline for H&E analysis exhibited high accuracy and consistency compared with five orthogonal data types within TRACERx, including DNA-seq, RNA-seq, IHC, 5,951 single-cell annotations by pathologists (balanced accuracy, as an average of specificity and sensitivity = 0.932), and pathology tumor-infiltrating lymphocyte (TIL) estimates following the guidelines developed by the International Immuno-Oncology Biomarker Working Group[7] (Extended Data Fig. 2, Supplementary Table 2). The Leicester Archival Thoracic Tumor Investigatory Cohort[8] (LATTICe-A, Extended Data Fig. 1c-d), a retrospective study of 970 resected LUAD patients that included H&E sections from all diagnostic tumor blocks with a median of four samples per tumor, was used for independent validation. The pipeline's generalizability was supported using 5,082 pathologists' single-cell annotations (balanced accuracy = 0.913), and virtual integration of IHC and H&E images generated from the same slides (Fig. 1e-h, Extended Data Fig. 2e-g, Supplementary Table 3). Using this unbiased scalable approach, immune infiltration was quantified as the percentage of all cells that were lymphocytes in each H&E image.

High geospatial immune variability between tumor regions within the same patients was revealed (Fig. 2a-b), which did not reflect associations with pathological stage (Extended Data Fig. 3). To differentiate highly from poorly immune infiltrated tumor regions, regions containing a lymphocyte percentage greater than a quarter standard deviation above the median lymphocyte percentage were classified as immune hot, and regions containing a lymphocyte percentage below a quarter standard deviation of the median were classified as immune cold. The remaining 20% were classified as intermediate (Fig. 2b). Subsequent results were tested on four more classification schemes based on the standard deviation to ensure that results derived from this classification were not contingent upon choice of thresholds used (Extended Data Fig. 4). Significant difference in pathology TIL estimates was observed between immune hot and cold regions (P = 4.6x10^{-8}, Extended Data Fig. 5a). Significantly higher levels of RNA-seq estimated immune infiltrate[1], particularly for immune

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

106

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

120

121 We investigated the impact of immune context on disease-free survival. Tumors with high
122 number of immune cold regions were at significantly increased risk of relapse that was
123 independent of the total number of regions sampled, tumor size and stage in both histology
124 types in TRACERx (Fig. 3d-e, Extended Data Fig. 6c-h). This association with disease-free
125 survival was also significant using the number of immune low regions as estimated by RNA-
126 seq[1] in 64 TRACERx tumors with available RNA-seq data (P = 0.002, Extended Data Fig. 6b).
127 Following the genomic findings in LUAD, we sought to validate this in 970 LUAD patients in
128 the multi-sample LATTICe-A cohort, confirming the prognostic value of immune cold sample
129 count, that was also independent of the number of samples per patient, tumor size and
130 stage (Fig. 3f-g, Extended Data Fig. 6c-e). In both cohorts, the number of immune cold
131 samples per patient correlated with relapse, more significantly than any other immune
132 feature generated using deep learning, including the average and variability of lymphocyte
133 percentage per tumor, number of immune hot regions, proportion of immune cold regions
134 to the number of regions sampled, as well as CD8$^+$ cell percentage or CD8$^+$ to CD4$^+$FOXP3$^+$
135 ratio in TRACERx diagnostic slides (Extended Data Fig. 6e).

136

137 Studies have revealed immunosuppressive fibroblast subsets localizing to the boundary of
138 tumor nests possibly contribute to T cell exclusion[11–13]. Therefore, we hypothesized that
139 increased cancer-stroma physical contact may reflect stroma-modulated inhibition of anti-
140 tumor immune responses[14–17]. To measure the physical contact between cancer and
141 stromal cells (the majority being fibroblasts) identified by image analysis, we developed a
142 spatial measure, using fractal dimension to quantify the geographical irregularity and
143 complexity of the cancer-stromal cell interface (Methods, Fig. 4a, Extended Data Fig. 7a,b,e).
144 Within the same tissue space, higher fractal dimension of cancer-stromal cell interface
145 suggests increased geometric irregularity and more extensive physical contact between
146 tumor and stromal cells than samples with a smooth interface. For both histology types,
147 fractal dimension was significantly higher in immune cold regions compared to immune hot
148 regions (Fig. 4b, Extended Data Fig. 7c). Moreover, the difference in fractal dimension
149 between immune cold and hot regions was more significant compared to the difference in
150 stromal cell percentage (both histology types combined: P = 0.00036, effect size 0.49 for
151 fractal dimension versus P = 0.018, effect size 0.38 for stromal cell percentage, Extended
152 Data Fig. 7d), suggesting the importance of stromal cell geographical location rather than
153 their quantity. This supports the hypothesis that the stroma-based inhibition of immune
154 infiltration[17] may result from a specific topological pattern in the form of cancer-stroma
155 engagement.
156
157 To understand the associations of stromal-mediated immunosuppression in the context of
158 the genetic mechanisms of immune evasion, we related fractal dimension to dysfunction in
159 antigen presentation through loss of heterozygosity at the human leukocyte antigen locus
160 (HLA LOH), which has been identified as a potent immune escape mechanism[1,18]. A
161 significantly higher fractal dimension was found in LUAD tumor regions with intact HLA
162 alleles compared with regions harboring HLA LOH (Fig. 4c, Extended Data Fig. 7f). This was
163 observed at the tumor level (see Methods for definition), independent of clonal neoantigen
164 burden (P = 0.04, multivariate regression, Extended Data Fig. 7h), but was not observed in
165 LUSC (Extended Data Fig. 7g, i).
166
167 Although clonal neoantigens have been associated with a cytotoxic immune response[19], the
168 spatial distribution of lymphocytes in relation to clonal neoantigens remained unclear. To
169 provide sufficient spatial context for analysis of cell distribution, whole-section TRACERx
170 diagnostic H&E images, typically 10x larger than the regional samples, were used. To test
171 the relationship between lymphocyte spatial distribution and clonal neoantigens, we
172 leveraged an established method for lymphocyte spatial modeling[20]. Each lymphocyte was
173 classified into three distinct spatial compartments: intra-tumor, adjacent-to-tumor or distal-
174 tumor, based on unsupervised modeling of cancer-lymphocyte proximity (Fig. 4d). In LUAD,
175 but not LUSC, clonal neoantigens[19] were found to be associated with a specific immune
176 spatial score to approximate pathology TIL estimates[7], defined as the ratio of adjacent-
177 tumor lymphocytes to stromal cells in the diagnostic H&E samples (P = 0.0074, high clonal

178   neoantigen defined as above median in LUAD, Fig. 4e; correlation as continuous variables
179   Rho = 0.37, P = 0.035 after multiple testing correction, Extended Data Fig. 8a). By contrast,
180   subclonal neoantigen burden did not correlate with any immune score (Extended Data Fig.
181   8a), supporting the notion that clonal but not subclonal neoantigens is associated with
182   infiltration of cytotoxic T cells[19] adjacent to tumor nests.

183

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

194

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

208

209   Spatial histology data can extend our knowledge of the tumor microenvironment
210   topological configuration in relation to genetic alterations relevant to immune surveillance,
211   including HLA LOH and clonal neoantigens in LUAD (Extended Data Fig. 9). Increased cancer-
212   stromal engagement as measured by fractal dimension may signal physical constraints
213   against T cell ingress. This is supported by previous studies in lung cancer showing
214   restriction of $CD8^+$ and $CD4^+$ T cell motility in dense stromal extracellular matrix areas
215   around tumor epithelial cell regions which prevent them from entering tumor islets[13].
216   Additionally, the association between specific spatial localization of lymphocytes in tumor-

217 adjacent stroma and clonal neoantigens further support exploration of the role of stromal
218 cells in limiting tumor infiltration by T cells[14–17].

219

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

Figures legends

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

261

262   **Figure 2. Geospatial heterogeneity of lymphocytic infiltration in the TRACERx cohort. a**.
263   Representative examples of immune hot and immune cold multi-region H&E samples, scale
264   bars represent 100μm. **b**. Each column represents a tumor, grouped by their histologic
265   subtype (the "Other" group consists of adenosquamous carcinoma, large cell
266   neuroendocrine carcinoma, pleomorphic carcinoma, and sarcomatoid carcinoma of
267   pleomorphic type arising from adenocarcinoma). Tumor regions (illustrated as dots) were
268   assigned to immune hot, immune cold, and intermediate phenotypes based on percentage
269   of lymphocytes in all cells following H&E-based deep learning analysis. CD8$^+$/CD4$^+$FOXP3$^-$
270   /CD4$^+$FOXP3$^+$ percentages based on automated analysis of the IHC diagnostic samples are
271   also shown. **c**. A heatmap showing gene expression patterns of 14 immune cell populations
272   across tumor regions, each row represents a tumor region (n = 142). The three clusters

273 correspond to the proposed immune regional classification as shown in **b**. **d.** Significant
274 enrichment of all immune cell populations in hot regions, as compared to cold regions,
275 particularly for the immune activating cell subsets, including cytotoxic, B-cell, and natural
276 killer cells (n = 109 regions; 52 patients). A two-sided, non-parametric, unpaired, Wilcoxon
277 signed-rank test was used for each box plot, all P-values were corrected for multiple
278 comparisons. Thick horizontal lines indicate the median value; outliers are indicated by the
279 extreme points; the first and third quantiles are represented by the box edges; and vertical
280 lines indicate the error range.

282 **Figure 3. Evolution of immune escape, and survival analysis in TRACERx and LATTICe-A. a**.
283 A box plot showing the difference in genomic distances for pairs of immune hot or immune
284 cold regions within the same patients in LUAD (n = 66 pairs). **b**. A box plot showing the
285 difference in mutational distance between the dominant subclones in pairs of immune hot
286 or immune cold regions via their last common ancestor in LUAD (n = 23 immune cold pairs;
287 15 immune hot pairs). This distance was calculated by taking the furthest dominant clone
288 (cancer cell fraction (CCF) ≥ 75%) from the trunk, and it remained significant when the
289 dominant clone closest to the most recent common ancestor of each tree was considered
290 (P = 0.02). **c**. Illustrative examples of tumor phylogenetic trees for a pair of immune hot and
291 immune cold regions. Dominant subclones were labelled and their last common ancestor
292 (annotated with arrows) was then identified. Minor (CCF < 75%) or undetected clones were
293 neglected in this analysis. **d,e**. Kaplan-Meier curves illustrating the difference in disease-free
294 survival according to the number of immune cold regions, dichotomized by the median
295 value, in TRACERx (**d**) (LUAD and LUSC, n = 79 patients, 249 regions) and LATTICe-A (**e**)
296 (LUAD, n = 970 patients, 4,324 samples). The same deep learning histology analysis and
297 immune regional classification developed for TRACERx were applied directly to LATTICe-A.
298 WSI: whole-section image. **f.** Forest plots showing multivariate Cox regression analyses in
299 TRACERx (n = 79 patients; LUAD and LUSC). Clonal neoantigens were dichotomized using the
300 upper quartile, determined individually for LUAD and LUSC tumors[1]. **g.** Forest plots showing
301 multivariate Cox regression analyses in LATTICe-A (n = 651 LUAD patients with complete
302 stage and smoking pack years data). For the patient subset with complete stage data but
303 missing pack years information, the test remained significant (n = 827, P < 0.001, HR =
304 1.4[1.1-1.9]). For statistical comparisons among groups, a two-sided, non-parametric,
305 unpaired, Wilcoxon signed-rank test was used, unless stated otherwise.

307 **Figure 4. Association of spatial histology with genetic alterations relevant to immune**
308 **surveillance. a.** An illustrative example of fractal dimension calculated by the box-counting
309 algorithm to quantify the geospatial complexity of the cancer cell-stromal cell interface. By
310 examining boxes of decreasing sizes that contain both cancer and stromal cells, the box
311 counting algorithm quantifies the rate at which the geometrical details of cancer-stromal
312 interface develop at increasingly fine scales. Blue box illustrates the smallest box of 20μm by
313 20μm in size. Scale bar represent 100μm. An example of a fractal structure displaying

314 geometrical self-similarity is shown below the panel. **b**. A box plot to illustrate the significant
315 difference in fractal dimension between all TRACERx immune hot and cold regions (n = 219).
316 **c**. A box plot showing a significant difference in fractal dimension between LUAD tumor
317 regions (n = 116) harboring an LOH event for class 1 HLA of any type versus regions that do
318 not, adjusted for multiple comparisons with the remaining HLA type-specific tests (see
319 Extended Data Fig. 7f). **d**. Illustration of the adjacent-tumor lymphocyte/stroma ratio
320 inferred by spatial modeling of cancer cell density (contours) and lymphocyte classification
321 into spatial compartments. Cell classification in IHC sample of the same block was shown for
322 comparison. Scale bars represent 50µm. **e.** A box plot showing the difference in the
323 adjacent-tumor lymphocyte/stroma ratio between high (≥ median) and low (< median)
324 clonal neoantigens for all LUAD patients in TRACERx (n = 61). **f**. Illustration of image
325 registration to spatially align serial sections of H&E and IHC and generate a virtual composite
326 map of T cell subset in the context of cancer/stroma density. T cell subsets classified in the
327 IHC were projected onto the cancer density map inferred from H&E, so that they can be
328 classified into adjacent-tumor, intra-tumor, and distal-tumor compartments. **g**. A box plot
329 showing **significantly higher** ratio of $CD8^+$ to $CD4^+FOXP3^+$ cells in adjacent-tumor and intra-
330 tumor lymphocytes compared with distal-tumor lymphocytes in registered LUAD image tiles
331 (n = 20 image tiles, using paired Wilcoxon test). For statistical comparisons among groups, a
332 two-sided, non-parametric, unpaired, Wilcoxon signed-rank test was used, unless stated
333 otherwise.

334

335

Main References

336 **Main References**
337

338 1. Rosenthal, R. et al. Neoantigen-directed immune escape in lung cancer evolution.
339 Nature 1 (2019). doi:10.1038/s41586-019-1032-7
340 2. Morris, L. G. T. & Chan, T. A. Lung Cancer Evolution: What's Immunity Got to Do with
341 It? Cancer Cell 35, 711–713 (2019).
342 3. Morris, L. G. T. et al. Pan-cancer analysis of intratumor heterogeneity as a prognostic
343 determinant of survival. Oncotarget 7, 10051–10063 (2016).
344 4. Milo, I. et al. The immune system profoundly restricts intratumor genetic
345 heterogeneity. Sci. Immunol. 3, (2018).
346 5. Jia, Q. et al. Local mutational diversity drives intratumoral immune heterogeneity in
347 non-small cell lung cancer. Nat. Commun. 9, (2018).
348 6. Jamal-Hanjani, M. et al. Tracking the Evolution of Non–Small-Cell Lung Cancer. N.
349 Engl. J. Med. 376, 2109–2121 (2017).
350 7. Hendry, S. et al. Assessing Tumor-Infiltrating Lymphocytes in Solid Tumors. Adv.
351 Anat. Pathol. 24, 311–335 (2017).
352 8. Moore, D. A. et al. In situ growth in early lung adenocarcinoma may represent
353 precursor growth or invasive clone outgrowth—a clinically relevant distinction. Mod. Pathol.
354 1 (2019). doi:10.1038/s41379-019-0257-1
355 9. Whittaker, K. A. & Rynearson, T. A. Evidence for environmental and ecological
356 selection in a microbe with no geographic limits to gene flow. Proc. Natl. Acad. Sci. U. S. A.
357 114, 2651–2656 (2017).
358 10. Shafer, A. B. A. & Wolf, J. B. W. Widespread evidence for incipient ecological
359 speciation: A meta-analysis of isolation-by-ecology. Ecology Letters 16, 940–950 (2013).
360 11. Costa, A. et al. Fibroblast Heterogeneity and Immunosuppressive Environment in
361 Human Breast Cancer. Cancer Cell 33, 463-479.e10 (2018).
362 12. Öhlund, D. et al. Distinct populations of inflammatory fibroblasts and myofibroblasts
363 in pancreatic cancer. J. Exp. Med. 214, 579–596 (2017).
364 13. Salmon, H. et al. Matrix architecture defines the preferential localization and
365 migration of T cells into the stroma of human lung tumors. J. Clin. Invest. 122, 899–910
366 (2012).
367 14. Thomas, D. A. & Massagué, J. TGF-β directly targets cytotoxic T cell functions during
368 tumor evasion of immune surveillance. Cancer Cell 8, 369–380 (2005).
369 15. Joyce, J. A. & Fearon, D. T. T cell exclusion, immune privilege, and the tumor
370 microenvironment. Science (80-.). 348, 74–80 (2015).
371 16. Sorokin, L. The impact of the extracellular matrix on inflammation. Nat. Rev.
372 Immunol. 10, 712–723 (2010).
373 17. Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer–immune set
374 point. Nature 541, 321–330 (2017).
375 18. McGranahan, N. et al. Allele-Specific HLA Loss and Immune Escape in Lung Cancer
376 Evolution. Cell 171, 1259-1271.e11 (2017).
377 19. McGranahan, N. et al. Clonal neoantigens elicit T cell immunoreactivity and
378 sensitivity to immune checkpoint blockade. Science (80-.). 351, (2016).
379 20. Yuan, Y. Modelling the spatial heterogeneity and molecular correlates of lymphocytic
380 infiltration in triple-negative breast cancer. J. R. Soc. Interface 12, 20141153 (2015).

381   21.   Thomas, A., Liu, S. V., Subramaniam, D. S. & Giaccone, G. Refining the treatment of
382   NSCLC according to histological and molecular subtypes. Nat. Rev. Clin. Oncol. 12, 511–526
383   (2015).
384   22.   Hammerman, P. S. et al. Comprehensive genomic characterization of squamous cell
385   lung cancers. Nature 489, 519–525 (2012).
386   23.   Collisson, E. A. et al. Comprehensive molecular profiling of lung adenocarcinoma:
387   The cancer genome atlas research network. Nature 511, 543–550 (2014).
388   24.   Keren, L. et al. A Structured Tumor-Immune Microenvironment in Triple Negative
389   Breast Cancer Revealed by Multiplexed Ion Beam Imaging. Cell 174, 1373-1387.e19 (2018).
390   25.   Giesen, C. et al. Highly multiplexed imaging of tumor tissues with subcellular
391   resolution by mass cytometry. Nat. Methods 11, 417–422 (2014).
392   26.   Goltsev, Y. et al. Deep Profiling of Mouse Splenic Architecture with CODEX
393   Multiplexed Imaging. Cell 174, 968-981.e15 (2018).
394

395

Author Contributions

K.A. and S.E.A.R. contributed equally to this work. S.E.A.R. and K.A. developed the image processing and deep learning pipeline and performed the geospatial analysis. K.A. performed the bioinformatics and statistical analyses. J.L.Q., R.S. and D.A.M. provided pathological expertise. M.J.-H. provided clinical expertise and patient characterization. S.V. performed histology sample generation and digitized H&E slides. A.A. generated and digitized IHC slides under the supervision of T.M. T.L. provided annotations for training and validating IHC analysis. N.M., R.R. and L.Z. assisted with genomic data integration. J.L.Q., R.S., S.L., M.A.B., D.A.M., C.T.H., and T.L. analyzed pathology TIL estimates. J.L.Q., L.O., M.S., and C. R. S. provided data and advice for LATTICe-A. Y.Y., N.M., J.L.Q., C.S., A.H. and S.A.Q. provided data analysis support and supervision. K.A., R.R., N.M., C.S. and Y.Y. wrote the manuscript with input from all authors. Y.Y. and C.S. jointly conceived and supervised the study.

Materials and Correspondence

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

Data availability

The digital pathology images from the TRACERx study generated or analysed during this study are not publicly available and restrictions apply to its use. A test subset of such digital pathology images are available through the Cancer Research UK & University College London Cancer Trials Centre (ctc.tracerx@ucl.ac.uk) for non-commercial research purposes and access will be granted upon review of a project proposal that will be evaluated by a TRACERx data access committee and entering into an appropriate data access agreement, subject to any applicable ethical approvals. Digital pathology images for LATTICe-A samples

478   with expert pathologist's annotations used for validation are available:

479   https://github.com/qalid7/compath. Request for data access for the remaining LATTICe-A

480   samples can be submitted to J.L.Q.

481

482   **Code availability**

483   The deep learning pipeline for digital pathology image analysis is available for non-

484   commercial research purposes: https://github.com/qalid7/compath. All code used for

485   statistical analyses of image data was developed in R version (3.5.1) and is available:

486   https://github.com/qalid7/tx100_compath.

487

Methods

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).

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 (n = 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.

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); anti-FOXP3 (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).

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 all patients, with complete pathological stage for 827 and smoking history for 651. All archival slides containing tumor material were used in order to capture the full diversity of each lesion. Slides were dearchived and scanned using a Hamamatsu NanoZoomer XR at 40x (226 nm/pixel resolution) yielding 15 TB of image data. Images containing incidental lymph

529 node tissue were excluded to avoid confounding immune infiltration analysis. For the
530 biological validation assay, a subset of 49 paraffin blocks from 49 patients was obtained
531 from the same study, and from these a validation TMA was prepared, containing a single
532 1mm core from each case. The work was ethically approved by an NHS research ethics
533 committee (ref. 14/EM/1159). This study complies with the STROBE guidelines.

534

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

549

550 For tissue segmentation, each whole slide image was reduced to 1.25x resolution and
551 segmented for tissue regions using Micro-Net-512[27] architecture. This architecture visualizes
552 the image at multiple resolutions, captures context information by connecting intermediate
553 deep layers and adds bypass connections to max-pooling to maintain weak features (Fig.
554 1b). 10 whole slide images were used to train the tissue segmentation network using Micro-
555 Net. The segmented images from the network were inspected visually and quantitatively
556 (Supplementary Table 6, Supplementary Figures 1-20) to evaluate performance using an
557 independent set of images.

558

559 The SCCNN adds two layers to conventional deep learning architecture for cell detection
560 within the segmented tissue. SC1 estimates the location and probability of each pixel
561 belonging to the center of a cell, and these probabilities are then mapped by SC2 to the
562 image. A customized implementation of SCCNN was coded in Python (version 3.5) using
563 TensorFlow[29] library (version 1.3) which makes it computationally more efficient compared
564 to the original MATLAB implementation[28]. To process an image of size 1000×1000 pixels, the
565 Python implementation takes 4.8 seconds for nucleus detection compared to 41.0 seconds
566 using the original implementation[28], excluding preprocessing which remained the same in
567 both implementations (using MATLAB (version 2018b)). In addition, through empirical
568 experimentation, we optimized the patch size to 31x31 instead of 27x27 in the original
569 implementation for increased cell detection accuracy. To generate nuclear locations from

the SC2 probability map, peak detection was applied where thresholds for intensity and minimum grouping distance were also optimized to 0.15 and 12 pixels through experimentation using validation data.

For cell classification, a neighboring ensemble predictor was used. This predictor utilizes SCCNN to classify cells in neighboring locations to the detected center of the cell. In our implementation, the ensemble classifier required votes from SCCNN classification of nine different neighborhood locations near to the center of the cell compared to five votes in original implementation. Through experimentation, the patch size was optimized to 51x51 for classification instead of 27x27 as originally proposed. This permitted incorporation of greater tissue spatial context while maintaining the accuracy of classifying small cells.

Altogether, this pipeline enabled the spatial mapping of four cell types from H&E images: cancer (malignant epithelial) cells, lymphocytes (including plasma cells), non-inflammatory stromal cells (fibroblasts and endothelial cells), and an "other" cell type that included non-identifiable cells, less abundant cells such as macrophages and chondrocytes, and 'normal' pneumocytes and bronchial epithelial cells.

Training the deep learning pipeline

To improve neural network generalizability and to avoid overfitting for cell detection and classification, we trained and tested our pipeline on a variety of sample types, including diagnostic (n = 100), regional (n = 275) and 133 cores corresponding to 75 TRACERx patients from TMA slides (63 patients had two cores and 12 patients had a single core). Both cell detection and classification were trained based on single-cell annotations from pathologists. Two thoracic pathologists annotated 26,960 cells on 53 whole slide images (3 TMAs, 35 regional slides and 15 diagnostic slides) to incorporate morphological variations in appearance of various cell types and stain variability. Several hundred examples of each cell class were marked on 76 cores selected at random from TMA images. In total, 4,056, 5,310, 15,007, 2,587 annotations were collected for stromal cells, lymphocytes, cancer cells and "other" cell types, respectively. These whole slide images were divided into small tile images of size 2000×2000 pixels (each pixel = 0.5µm), which were then divided into three sample sets maintaining the class distribution of cells. These included: 13 diagnostic, 58 regional and 134 TMA tile images for training; 4 diagnostic, 21 regional and 72 TMA tile images for validation; and 3 diagnostic, 22 regional and 61 TMA tile images for testing. As a result, the annotations were divided between the three groups; 2/3 for training, 1/6 for validation and 1/6 for testing. The training set included annotations for 2,147 stromal cells, 3,183 lymphocytes, 10,103 cancer and 1,357 other cell types. The validation set had annotations for 473 stromal cells, 825 lymphocytes, 2,562 tumor and 359 other cell types. Breakdown for the test set is provided in Supplementary Table 2.

610   For IHC cell classification, we used a pretrained SCCNN network on samples stained for
611   CD4/CD8/FOXP3. The training set consisted of 1,657 CD4[+]FOXP3[-], 3,187 CD8[+], 1,001
612   CD4[+]FOXP3[+], and 3,488 other (negative) cells. The trained network was tested on 5,028 cell
613   annotations collected on 6 lung diagnostic whole slide images, including 251 CD4[+]FOXP3[-],
614   406 CD8[+], 123 CD4[+]FOXP3[+] and 4,248 other cells to test the ability of the algorithm in
615   correctly detecting and classifying negative cells. See Supplementary Table 7 for the total
616   number of identified cells in the H&E diagnostic, H&E multi-region and IHC diagnostic
617   datasets.
618

619 **Validation of the H&E deep learning pipeline with orthogonal data types**
620   The algorithms' performance in detecting and classifying single cells in H&E were first
621   evaluated against the test set of 5951 cells. Individual class accuracy statistics were
622   calculated using the R function 'confusionMatrix' from the R package 'caret'.
623

624   Pathology TIL estimates were scored following the international guidelines developed by the
625   International Immuno-Oncology Biomarker Working Group[7]. Briefly, by inspection of H&E
626   slide of a given tumor region, the fraction of the stromal area infiltrated by TILs was
627   assessed.
628

629   For regional samples, tumor cellularity, estimated as the computed percentage cancer cells
630   was correlated with tumor purity estimated by ASCAT based on DNA-seq copy number and
631   VAF purity (both available from Jamal-Hanjani et al.[6], n = 239 regional tumor samples). The
632   RNA-seq-based CD8[+] T cell signature (available from Rosenthal et al.[1], computed using the
633   Danaher et al. method[30]) was correlated with the deep learning based lymphocyte
634   percentage for 142 regional tumor samples. For diagnostic samples, deep learning-based
635   lymphocyte percentage from H&E was correlated with deep learning-based CD8[+] cell
636   percentage from IHC (n = 100 diagnostic samples, Extended Data Fig. 2a-d).
637

638   Discordance rate between RNA-seq based[1] and histology/deep learning-based immune hot
639   and cold regional classification was calculated by cross-tabulation of immune hot and cold
640   (from histology) versus high and low (from RNA-seq), disregarding any regions without one
641   of these two types of data. The RNA-seq method used 15 immune cell signatures presenting
642   different T- and B-cell subsets, as well as neutrophils, macrophages, mast and dendritic cells,
643   to classify tumor regions into high and low categories. A Fisher's exact test was used to
644   compute the overlap between the two immune classifications. Distributions of multiple
645   immune scores (lymphocyte percentage, intra-tumor lymphocytes and adjacent-tumor
646   lymphocytes/stroma) as well as ASCAT tumor purity were compared between hot versus
647   cold (deep learning) and high versus low (RNA-seq) classifications (Extended Data Fig. 5).
648

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

650 The external validity of the proposed deep learning pipeline was performed on 100
651 randomly selected patients from the LATTICe-A cohort[8]. This validation ensures that the
652 trained cell detection and cell classification models from the TRACERx tumor blocks are
653 generalizable to a distinct dataset which is processed, stained and scanned in another
654 center (the LATTICe-A study, University of Leicester).
655
656 All 100 whole-tumor H&E sections were processed using the same TRACERx trained model.
657 The validation was then performed using two data types. First, a pathologist provided 5,082
658 single-cell annotations following the same protocol for TRACERx in 20 randomly selected
659 LATTICe-A sections. The breakdown for single-cell annotations was 1,997 stromal cells, 787
660 lymphocyte cells, 1,839 cancer cells and 459 other cells (see Supplementary Table 3).
661 Second, two independent pathologists jointly scored the remaining 80 sections for overall
662 fraction of lymphocytic infiltration and pathology TIL estimates[7]. These manual scores were
663 correlated with the deep learning-based lymphocyte percentage and adjacent-tumor
664 lymphocytes/total stroma (Extended Data Fig. 2e).
665
666 **Validation of the deep learning pipeline with biological assays**
667 A new biological validation method was developed to overcome the challenge of obtaining
668 large quantities of cell-specific validation data (Fig. 1f-h, Extended Data Fig. 2f-g). 48 cores
669 were available for the TTF1-H&E image pairs, 38 for the CD45-H&E pairs, and 33 for the
670 SMA-H&E pairs. Stains were performed using a Ventana BenchMark ULTRA instrument
671 (H&E, TTF-1) or a Dako Link 48 (CD-45, SMA). Digital images were acquired using a
672 Hamamatsu Nanozoomer slide scanner. First, H&E staining was performed using a Leica
673 Infinity kit, and a digital image was collected. The slide was subsequently de-coverslipped,
674 the H&E stain removed by acid alcohol washing, and then an immunohistochemical stain
675 with haematoxylin counterstain was applied using a standard diagnostic antigen retrieval
676 and antibody protocol. A second digital image was acquired after mounting and
677 coverslipping. Through experimentation, no difference in the staining was observed when
678 the procedure was reversed.
679
680 TTF-1 (type: Novocastra Liquid Mouse Monoclonal antibody thyroid transcription factor 1,
681 clone: SPT24, cat. no.: NCL-L-TTF-1, source: Leica biosystems, Germany, used at 1:100) was
682 selected as the cancer cell marker in these LUAD samples because it is the most robust and
683 widely used immunohistochemical marker of LUAD cells[31]. It is very specific, both in that
684 only epithelial cells are stained in the lung, and in that very few tumors of non-lung or
685 thyroid origin are stained[32]. The sensitivity of the antibody clone used (SPT24) is also high,
686 staining >75% of tumor cells in 76% of LUAD tumors in one published series[33]. However, as
687 this implies, there are many tumors in which tumor cell staining is incomplete (i.e. <100%).
688 Therefore, only cores showing near-universal TTF-1-positivity of tumor cells were used for
689 validation, in order to provide the best possible 'gold standard' comparator for the deep
690 learning algorithm. The same procedure was followed for pairs of H&E-CD45 (anti-human

691 CD45, type: Mouse Monoclonal, clone: 2B11 + PD7/26, cat. no.: M0701, source: Agilent
692 DAKO, USA, used at 1:200) and H&E-SMA (myofibroblast marker, type: Mouse Monoclonal
693 antibody Smooth Muscle Actin (1A4), cat. no.: 760-2833, source: Roche, Switzerland, a
694 ready to use antibody) to biologically validate the accuracy of single cell classification.
695
696 In total, 64,976 TTF1$^+$ cells, 26,284 CD45$^+$ cells and 46,343 SMA$^+$ cells were detected from
697 the IHC images, denoting the advantage of this method in acquiring large amount of
698 validation data at single-cell resolution. The correlation measured (Fig. 1f-h, Extended Data
699 Fig. 2g) was that between the fraction of classified cells in the H&E versus fraction of
700 positively stained IHC cells per 100μm^2.
701

Immune phenotype classification
703 To classify tumor regions into different immune phenotypes, we assigned each region to an
704 immune hot, cold or intermediate category based on lymphocyte percentage. The
705 dependency of our subsequently results on thresholds chosen for this classification scheme
706 was tested after applying perturbations to the thresholds used. Four new classification
707 schemes were tested: no intermediate zone (i.e. using median lymphocyte percentage for
708 separating hot and cold regions), regions with lymphocyte percentage greater than standard
709 deviation/2 above/below the median lymphocyte percentage classified as immune hot/cold,
710 , and similarly for standard deviation/3 and standard deviation/6 (Extended Data Fig. 4a-b).
711 For every new classification, we repeated the multivariate survival analysis to confirm the
712 significance of the number of immune cold regions in predicting disease-free survival as well
713 as the genomic distance test for pairs of immune hot versus immune cold regions in LUAD
714 patients (Extended Data Fig. 4b). In addition, the CD8$^+$ RNA-seq signature was used to test
715 the difference in CD8$^+$ levels between immune hot and immune cold phenotypes across all
716 classification schemes (Extended Data Fig. 4c).
717

Genomic distance measure
719 Genomic distance was calculated as described previously[1], by taking the Euclidean distance
720 of the mutations present for every pair of immune hot and immune cold regions from the
721 same patient. All mutations present in a region from a tumor were turned into a binary
722 matrix of which the rows were mutations and columns were the tumor regions. From this
723 matrix, the pairwise distance was determined.
724

Distance between dominant clones to the last common ancestor of region pair
726 Deep learning-based immune phenotypes were integrated with the TRACERx phylogenetics
727 data[6]. Dominant clones (using the upper quartile of cancer cell fraction, ≥ 75%) were
728 labelled for all tumor regions' trees which had an available H&E sample in LUAD patients (n
729 = 76 regions, 15 immune hot pairs and 23 immune cold pairs). For every pair of immune hot
730 / cold regions within a tumor, the distance between the dominant clones (as measured by
731 branch length, i.e. number of mutations) via their last common ancestor was computed. The

732   recently shared ancestry clone between the two dominant clones was labelled as the 'last
733   common ancestor of region pair' (annotated with arrows in Fig 3.c). To ensure this analysis
734   was not dependent on a certain cancer cell fraction threshold, multiple thresholds (CCF ≥
735   80%, 85%) were placed while repeating the same analysis. Next, by identifying the last
736   common ancestral subclone for pairs of the same phenotype, each pair was categorized into
737   one of two diversification patterns: 'diversifying at the most recent common ancestor
738   (MRCA) of the tree' or 'diversifying at a descendant subclone of the MRCA of the tree'. The
739   latter category included a pattern exclusive to immune cold pairs, where the two regions
740   shared the same dominant subclone that was the direct descendant of the MRCA of the
741   tree.
742

743   **Tumor spatial modelling**
744   H&E and IHC cell abundance scores (e.g. lymphocyte percentage, CD8$^+$ percentage) were
745   computed as the percentage of a cell type in the total sample cell count. Stromal TILs were
746   identified using spatial modelling[20,34,35], where lymphocytes were classified (using
747   unsupervised clustering) into intra-tumor lymphocytes, adjacent-tumor lymphocytes and
748   distal-tumor lymphocytes based on their spatial proximity to epithelial cell nests in H&Es.
749   The immune hotspot score was calculated using the Getis–Ord algorithm as previously
750   described[36]. To capture the emergence of complex morphological patterns that dictate
751   cancer-stromal cell spatial contact preserved over varying spatial scales, a fractal dimension
752   calculation (Minkowski-Bouligand dimension) was performed using the box-counting
753   algorithm[37]. This algorithm calculates the number of boxes of a certain size needed to cover
754   a geometric pattern. We modified a MATLAB-based algorithm[38] to include both spatial
755   information of cancer and stromal cells, as opposed to its conventional use on one variable
756   (i.e. pixel information of an image). The analysis was carried out on spatial maps generated
757   using coordinates of classified stromal and cancer cells, while utilizing the tissue segmented
758   image (as a boundary mask) to exclude all empty tissue areas. Choices of box size were
759   informed by the distribution of minimum and maximum Euclidean distance for each stromal
760   cell to its nearest cancer cell in all 275 tumor regions (Extended Data Fig. 7a). The mean
761   minimum distance was 21.43µm. We limited the upper box size at 300µm, which is just
762   above a previously proposed cell-cell communication distance of 250µm[39] but designed to
763   be more inclusive. For statistical tests where fractal dimension was represented at tumor
764   level, the maximum regional score was used.
765

766   **H&E-IHC spatial alignment/immune subset projection**
767   For a H&E diagnostic slide, we determined the number of intra-tumor lymphocytes,
768   adjacent-tumor lymphocytes and distal-tumor lymphocytes (n_I, n_A, n_D) based on spatial
769   modelling of the H&Es. After spatial alignment of IHC and projecting IHC-derived cells onto
770   the H&E, the number of CD8$^+$ cells that were also intra-tumor lymphocytes was determined
771   (n^{CD8}_{ITL}), and similarly for other cell types. As a result, intra-tumor lymphocytes were
772   deconvoluted by $n_I = n^{CD8}_I + n^{CD4}_I + n^{FOXP3}_I + n^{other}_I$. Two-sided paired Wilcox was used to test the

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

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

775   for CD4$^+$FOXP3$^-$ and CD4$^+$FOXP3$^+$ cells.

776

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

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

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

780   analysis was performed on the remaining nine samples. The quality of alignment was

781   evaluated by manually identifying 238 visible landmarks and placed on corresponding

782   positions in H&E and IHC tiles (total number of tiles = 249, maximum landmarks per tile = 5),

783   as shown in Extended Data Fig. 8b. These marked points were used to compute the

784   Euclidean distance (difference in x, y coordinates) between them to obtain a quantitative

785   measurement of alignment accuracy. The average distance between matching landmarks

786   was 9.57µm, whereas the maximum distance between the H&E and CD4/CD8/FOXP3

787   sections was 16µm.

788

789   **Survival analysis and other statistical methods**

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

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

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

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

794   parameters included in the multivariate model were age, sex, smoking pack years, histology

795   (whether LUAD, LUSC or otherwise), tumor stage, adjuvant therapy (whether received or

796   not). Because of its prognostic importance in TRACERx, the upper quartile of clonal

797   neoantigens in each histology cohort was also incorporated in the multivariate model. The

798   range of available disease-free survival data was 34-1364 days (median = 915 days) in

799   TRACERx, and 1-6139 days (median = 684 days) in LATTICe-A. All hazard ratios were

800   computed on all time points (i.e. the whole survival curve, not at a specific time point).

801   Correlation tests used Spearman's method and were generated using the function

802   'ggscatter' from the 'ggpubr' R package. All correlation plots show the Rho (ρ) coefficient

803   and the significance P-value. For statistical comparisons among groups, a two-sided, non-

804   parametric, unpaired, Wilcoxon signed-rank test was used, unless stated otherwise. All box

805   plots were generated using the function 'ggboxplot' from the 'ggpubr' R package (all data

806   points are plotted with the 'jitter' option, the median value is indicated by a thick horizontal

807   line; minimum and maximum values are indicated by the extreme points; the first and third

808   quantiles are represented by the box edges; and vertical lines indicate the error range) or

809   the function 'ggbetweenstats' from the 'ggstatplot' R package for more than two groups.

810   Tests for concordance between two data classes were analyzed using a Fisher's exact test.

811   All statistical tests were two-sided, a P value of less than .05 was considered statistically

812   significant. To adjust P-values for multiple comparisons, the Benjamini & Hochberg method

813   was used. To measure effect size, Cohen's d method was used. All statistical analyses were

814   conducted in R (version 3.5.1).

815

816   **Reporting summary**

817   Further information on research design is available in the Nature Research Reporting

818   Summary linked to this paper.

819

Extended Data Figures legends

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

831

832 **Extended Data Fig. 2. Validation of the automated single-cell classification for H&E. a.** A
833 scatter plot showing the correlation between H&E-based adjacent-tumor
834 lymphocytes/stromal and pathology TIL estimates in diagnostic samples (n = 98 diagnostic
835 slides/patients). **b.** Scatter plots showing the correlations between H&E-based tumor
836 cellularity estimate and ASCAT/VAF purity scores (n = 238 regions; 83 patients). **c.** A scatter
837 plot showing the correlation between H&E-based estimate of lymphocyte percentage
838 among all cells and RNA-seq-based CD8$^+$ signature using the Danaher et al. method[30] (n =
839 142 regions; 56 patients). **d.** A scatter plot showing the correlation between H&E-based
840 estimate of lymphocyte percentage among all cells and CD8$^+$ cell percentage in IHC in the
841 diagnostic samples (n = 100 diagnostic slide/patients). **e.** Scatter plots showing the
842 correlation between H&E-based lymphocyte percentage versus pathological scores of
843 overall lymphocytic cell fraction, and adjacent-tumor lymphocytes/stromal versus pathology
844 TIL estimates in an external cohort (LATTICe-A, n = 80 diagnostic slides/patients). **f.**
845 Illustrative example to show the spatial alignment of TTF1/CD45/SMA-stained IHC and H&E
846 images obtained using sequential staining on the same tissue microarray section for
847 biological validation. **g.** A scatter plot showing the correlation between stromal cell
848 percentage determined by H&E and SMA$^+$ cell percentage per LUAD image tiles of size
849 100μm^2 (n = 144). The experiment was conducted once using one TMA (n = 33
850 cores/patients). The shading indicates 95% confidence interval.

851

852 **Extended Data Fig. 3. Distribution of regional lymphocytic infiltration according to**
853 **pathological stage.** All available patients' data have been used in this figure except for the
854 standard deviation tests excluding patients with a single tumor region. Patients without
855 pathological staging information from the LATTICe-A cohort were also removed. **a, b, c**, top
856 row: TRACERx and bottom row: LATTICe-A. Horizontal lines indicate the median value. **a.**
857 Distribution of the standard deviation of regional lymphocyte percentage for LUAD and
858 LUSC patients in TRACERx (n = 69), and LUAD in LATTICe-A (n = 814). **b.** Distribution of the

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

870

871   **Extended Data Fig. 4. Validation of immune phenotype classification. a.** The proposed
872   immune classification imposed on density plot showing distribution of lymphocyte
873   percentage. The middle zone corresponds to the intermediate phenotype, red zone for
874   immune hot and blue zone for immune cold. Black dash line shows the median. This
875   classification was validated after applying small perturbations to the thresholds to re-classify
876   regional immune phenotypes, illustrated as grey dash lines: no intermediate zone (i.e. hard
877   median for separating hot and cold), standard deviation (SD)/2 above and below the
878   median, SD/3 and SD/6. **b**. Forest plots to show repeated multivariate Cox regression tests
879   for the number of immune cold regions using these new classifications (n = 79 patients),
880   after accounting for stage, total number of samples, upper quartile of clonal neoantigens
881   determined for LUAD and LUSC individually, and other clinical parameters. Box plots
882   showing difference in genomic distance for pairs of hot regions compared with pairs of cold
883   regions for LUAD and LUSC separately (LUAD: n = 45 hot pairs, 45 cold pairs for no
884   intermediate zone; n = 19 hot, 25 cold for SD/2; n = 25 hot, 33 cold for SD/3; n = 32 hot, 41
885   cold for SD/6. LUSC: n = 32 hot pairs, 54 cold pairs for no intermediate zone; n = 19 hot, 27
886   cold for SD/2; n = 19 hot, 37 cold for SD/3; n = 27 hot, 41 cold for SD/6.). **c**. Box plots
887   showing significant difference in CD8$^+$ RNA-seq signature using the Danaher method
888   between regions of hot and cold phenotype across all classification schemes (n = 219 for
889   SD/4; 275 for no intermediate zone; 173 for SD/2; 204 for SD/3; 237 for SD/6). **d**.
890   Distribution and difference of lymphocytic infiltration for LUAD versus LUSC regions in
891   TRACERx (n = 275 regions; 85 patients) as well as distribution for LUAD in LATTICe-A (n =
892   4,324 samples; 970 patients). Horizontal lines in the distribution plots indicate mean values.
893   For statistical comparisons among groups, a two-sided, non-parametric, unpaired, Wilcoxon
894   signed-rank test was used, unless stated otherwise.

895

896   **Extended Data Fig. 5. Concordance between histology deep learning and RNA-seq immune
897   classification. a**. A box plot showing the difference in pathology TIL estimates between

898   immune hot and immune cold regions (n = 219). Pathology TIL estimates score fraction of
899   stroma containing TILs, whereas immune classification was defined based on the percentage
900   of lymphocytes in all cells within a slide. **b**. A confusion matrix to compare RNA-seq and
901   deep learning histology immune classifications (discarding immune intermediate regions, n
902   = 109 regions (57 LUAD, 37 LUSC, 15 other histology subtypes); 52 patients). The p-value
903   was generated using a two-sided Fisher's exact test for overlap. **c**. A box plot showing the
904   difference in the fraction of immune hotspots[36] in regions where the two classifications are
905   in agreement (n = 78; labeled as 'In agreement') against the discrepant regions (n = 31,
906   labeled as 'Discrepant'). Each dot represents a region, the median value is indicated by a
907   thick horizontal line; minimum and maximum values are indicated by the extreme points;
908   and the first and third quantiles are represented by the box edges. **d**. Box plots to support
909   the overall consistency between H&E-deep learning and RNA-seq methods by comparing
910   different immune scores as well as ASCAT tumor purity between immune hot/high and
911   cold/low tumor regions (all P-values < 0.0001). Top row, H&E-deep learning immune
912   classification (n = 219; except the ASCAT purity box plot n = 186 regions), bottom row, RNA-
913   seq derived immune classification (n = 142; except the ASCAT purity box plot, n = 141
914   regions). For statistical comparisons among groups, a two-sided, non-parametric, unpaired,
915   Wilcoxon signed-rank test was used, unless stated otherwise.

916

917   **Extended Data Fig. 6. Genomic and survival analysis of tumor regions according to**
918   **immune phenotypes. a.** A box plot showing the difference in genomic distances for pairs of
919   immune hot versus immune cold regions within the same LUSC patients (n = 59 pairs). A
920   two-sided, non-parametric, unpaired, Wilcoxon signed-rank test was used. **b.** Forest plots to
921   show the univariate prognostic value for the number of immune low regions (both as
922   continuous and dichotomized at the median (≤1 versus >1)), or the number of immune high
923   regions, using the immune classification generated by RNA-seq-based infiltrating immune
924   cell populations[1] in 64 TRACERx tumors (41 LUAD, 16 LUSC and 7 other histology subtypes)**.**
925   **c.** Forest plots showing multivariate Cox regression analyses in both TRACERx (n = 79
926   patients; LUAD and LUSC combined) and LATTICe-A (n = 651 LUAD patients representing a
927   subset with complete stage and smoking pack years data) with the number of immune cold
928   regions dichotomized at the median (≤1 versus >1). This remains significant when the
929   number of immune cold regions was replaced as a continuous variable, in the same
930   multivariate model, (P = 0.019 in TRACERx and < 0.001 in LATTICe-A, for the number of
931   immune cold regions). Clonal neoantigens were dichotomized using the upper quartile,
932   determined individually for LUAD and LUSC tumors[1]. **d**. The same test in **c** when tumor size
933   (in mm) was also controlled in the multivariate model in LATTICe-A. This test also remained
934   significant for a bigger group of patients with complete stage data, but missing pack years
935   information (n = 815, P< 0.001, HR = 1.4[1.1-1.8]). **e**. Forest plots to compare the prognostic
936   value of regional immune scores as well as diagnostic H&E and IHC scores for relapse-free
937   survival in TRACERx (n = 79 patients, LUAD and LUSC combined). Wherever possible, these

938 immune features were tested in LATTICe-A (n = 970 patients). To compare the prognostic
939 value of the number of immune cold region with other immune features, LATTICe-A
940 comparisons were conducted in Cox multivariate regression models to include every
941 immune feature after correcting for the number of immune cold regions in the same model.
942 Each variable's HR is plotted with a 95% confidence interval; all P-values were adjusted for
943 multiple testing; and the size of the circles denotes $-\log_{10}(P)$. For the sake of visualization,
944 a minor adjustment was made to the HR for the number of cold regions/total number of
945 regions in LATTICe-A from 0.88[0.57-1.3] to 0.99[0.97-1.3]. SD: standard deviation, used for
946 measuring variability of lymphocyte percentage among samples within a tumor. **f.** Forest
947 plots using Cox multivariate regression analysis showing that the prognostic value of the
948 number of immune cold regions was independent of: 1) genetic measure, subclonal copy
949 number alteration (obtained from [6]); 2) tumor cellularity from DNA-seq-based ASCAT purity,
950 3) tumor cellularity measured by deep learning-based cancer cell percentage. **g.** Kaplan
951 Meier curves to illustrate the difference in relapse-free survival for TRACERx patients
952 including other histology types (n = 85; representing all TRACERx patients in the multi-
953 region histology cohort) with high and low number of immune cold regions, dichotomized
954 by its median value. Log-rank P= 0.0017. **h.** Forest plot using Cox regression for the
955 multivariate survival analysis for the number of immune cold regions in TRACERx including
956 patients with other histology subtypes (n = 85).

957

958 **Extended Data Fig. 7. Fractal dimension and relationships with stromal cells. a.**
959 Distribution of the average minimum Euclidean distance between a stromal cell to its
960 neighboring cancer cell. For every stromal cell in a tumor region slide, the minimum distance
961 to nearest cancer cell was computed. This distance was then averaged for all identified
962 stromal cells in every region to plot the distribution (n = 275 regions; 85 patients). **b**.
963 Distribution of the fractal dimension of the cancer-stroma cell interface for histology types
964 in the TRACERx cohort (n = 275 regions; 85 patients). **c.** Box plots to show the difference in
965 fractal dimension between immune hot and cold regions in TRACERx LUAD (n = 113) and
966 LUSC (n = 84). **d**. Box plots showing the difference in stromal cell percentage between
967 immune hot and cold regions in all (n = 219), LUAD (n = 113), and LUSC (n = 84). **e**. Scatter
968 plots showing the correlation between fractal dimension and percentage of cells that are
969 stromal or cancer in all tumor regions (n = 275 regions; 85 patients). This shows that fractal
970 dimension was independent of tumor cell composition, with only a weak correlation with
971 stromal cell percentage and no correlation with tumor cellularity. **f**. Box plots showing the
972 difference in fractal dimension between LUAD tumor regions harboring an LOH event for
973 HLA type A (n = 106), type B (n = 113), type C (n = 108) versus regions that do not, adjusted
974 for multiple comparisons with the corresponding test in Fig. 4c. **g**. The same test in **f**
975 repeated for LUSC tumor regions (n = 87) for HLA of any type. **h.** Box plots showing the
976 difference in tumor-level fractal dimension using the maximum value of regional measures
977 between LUAD tumors (n = 48) harboring a single LOH event for any HLA type, HLA type A,

978　type B and type C versus tumors that do not, independent of predicted clonal neoantigens.
979　Each p-value was generated using a multiple regression linear model and was also adjusted
980　for multiple testing correction. **i**. The same test in **h** repeated for LUSC tumors (n = 29) for
981　HLA of any type. For statistical comparisons among groups, a two-sided, non-parametric,
982　unpaired, Wilcoxon signed-rank test was used, unless stated otherwise.

983

984　**Extended Data Fig. 8. Relationship of immune subsets and spatial TILs in LUAD. a.**
985　Spearman's correlations between immune scores in diagnostic slides and genetic measures
986　including predicted neoantigens and HLALOH in LUAD patients (n = 46). ITLR: intra-tumor
987　lymphocytes to total tumor cell ratio. Only significant correlations after multiple testing are
988　highlighted (rho = 0.37, P = 0.035). **b**. Examples of registered H&E and IHC tiles. The green
989　cross denotes a manually placed landmark repeated 238 times on pairs of H&E-IHC image
990　tiles. The Euclidean distance (difference in x, y coordinates) was computed between the two
991　landmarks which was then **c.** shown as a distribution to represent the accuracy of the
992　registration (n = 249 total H&E-IHC image tiles, maximum five landmarks per a pair of tiles).
993　The average distance between matching landmarks was 9.57µm and the distribution is
994　within the expected range of maximum distance between four serial sections (16µm). **d.** Box
995　plots to illustrate the difference in percentage of immune cell subsets among adjacent, intra
996　and distal-tumor lymphocytes (n = 20 image tiles), a non-parametric, paired Wilcoxon test
997　was used.

998

999　**Extended Data Fig. 9. Summary of immune and genomics features in NSCLC.** An extended
1000　heatmap showing all immune variables described in TRACERx across all patients (n = 275
1001　regions; 85 patients), along with genetic measures and clinical parameters. Each column
1002　represents a tumor, grouped by their histologic subtype. Tumor regions (illustrated as dots)
1003　were assigned to immune hot, immune cold and intermediate phenotypes based on
1004　percentage of lymphocytes in all cells following H&E-based deep learning analysis. Cancer-
1005　stromal fractal dimension, defined using the maximum fractal dimension in regions of a
1006　patient, using the median as cut-off to determine high and low groups.

1007

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).

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

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

20. Yuan, Y. Modelling the spatial heterogeneity and molecular correlates of lymphocytic infiltration in triple-negative breast cancer. *J. R. Soc. Interface* **12**, 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: The
1098   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 in
1107   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**, 49–
1117   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 on
1122   Small Biopsies or Cellblocks for Typing Non-small Cell Lung Cancer: A Novel Two-Hit,
1123   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 of
1133   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 Dubois-Marshall (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 Papadatos-Pastos (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

(6)	Department of Cellular Pathology, University College London, University College Hospital, London, UK

(7)	Translational Immune Oncology Group, Centre for Molecular Medicine, Royal Marsden Hospital NHS Trust, London, UK

(8)	Department of Pathology, GZA-ZNA-Ziekenhuizen, Antwerp, Belgium

(9)	Division of Research, Peter MacCallum Cancer Centre, University of Melbourne, Melbourne, Victoria, Australia

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

1220 (11)   Leicester Cancer Research Centre, University of Leicester, Leicester, UK

1221 (12)   Cancer Research UK & University College London Cancer Trials Centre, University College
1222 London, London, UK

1223 (13)   Cancer Immunology Unit, University College London Cancer Institute, London, UK

1224 (14)   Cancer Genome Evolution Research Group, University College London Cancer Institute,
1225 University College London, London, UK

1226 (15)   Glenfield Hospital, University Hospitals Leicester NHS Trust, Groby Road, Leicester, UK

1227 (16)   Cancer Research Centre, University of Leicester, Leicester, United Kingdom

1228 (17)   The Francis Crick Institute, London, United Kingdom

1229 (18)   University College London Cancer Institute, London, United Kingdom

1230 (19)   Aberdeen Royal Infirmary, Aberdeen, United Kingdom

1231 (20)   Ashford and St Peter's Hospitals NHS Foundation Trust

1232 (21)   Barnet & Chase Farm Hospitals, United Kingdom

1233 (22)   Barts Health NHS Trust

1234 (23)   Berlin Institute for Medical Systems Biology, Max Delbrueck Center for Molecular Medicine,
1235 Berlin, Germany

1236 (24)   German Cancer Consortium (DKTK), partner site Berlin

1237 (25)   German Cancer Research Center (DKFZ), Heidelberg

1238 (26)   Cancer Research UK Manchester Institute, University of Manchester, Manchester, UK

1239 (27)   Cancer Research UK Lung Cancer Centre of Excellence, University of Manchester,
1240 Manchester, UK

1241 (28)   Leicester University Hospitals, Leicester, United Kingdom

1242 (29)   Cardiff & Vale University Health Board, Cardiff, Wales

1243 (30)   Christie NHS Foundation Trust, Manchester, United Kingdom

1244 (31)   Danish Cancer Society Research Center, Copenhagen, Denmark

1245 (32)   Department of Physics of Complex Systems, ELTE Eötvös Loránd University, Budapest,
1246 Hungary

1247 (33)   Departments of Radiation Oncology and Radiology, Dana Farber Cancer Institute, Brigham
1248 and Women's Hospital, Harvard Medical School, Boston, MA, USA

1249 (34)   Department of Radiology, Netherlands Cancer Institute, Amsterdam, The Netherlands

1250 (35)   Golden Jubilee National Hospital

1251 (36)   Independent Cancer Patients Voice

1252 (37)   University of Leicester, Leicester, United Kingdom

1253 (38) Liverpool Heart and Chest Hospital NHS Foundation Trust

1254 (39) Royal Liverpool University Hospital

1255 (40) Manchester Cancer Research Centre Biobank, Manchester, United Kingdom

1256 (41) National Institute for Health Research Leicester Respiratory Biomedical Research Unit,
1257 Leicester, United Kingdom

1258 (42) NHS Greater Glasgow and Clyde

1259 (43) Royal Brompton and Harefield NHS Foundation Trust

1260 (44) Sheffield Teaching Hospitals NHS Foundation Trust

1261 (45) The Princess Alexandra Hospital NHS Trust

1262 (46) The Whittington Hospital NHS Trust, London, United Kingdom

1263 (47) University College London Hospitals, London, United Kingdom

1264 (48) University Hospital Birmingham NHS Foundation Trust, Birmingham, United Kingdom

1265 (49) University Hospital Southampton NHS Foundation Trust

1266 (50) Velindre Cancer Centre, Cardiff, Wales

1267 (51) Wythenshawe Hospital, Manchester University NHS Foundation Trust

1268 (52) Division of Infection, Immunity and Respiratory Medicine, University of Manchester,
1269 Manchester, UK

1270

a TRACERx multiregion sampling

Diagnostic block

R1 R3 R4 R2

DNA/RNA ← → FFPE/H&E

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

c H&E

Cancer

Stromal

Lymphocyte

Other

d IHC

CD8+

CD4+FOXP3−

CD4+FOXP3+

Other

Other   CD4+FOXP3−

CD4+FOXP3+

e Annotated data (TRACERx)

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

Deep learning model

TRACERx → 275 multiregion WSIs, 100 diagnostic WSIs, 100 patients

LATTICe-A → 4,324 WSIs, 970 patients

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

5 Billion single-cells identified

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

H&E−based cancer cell / 100μm²

h rho = 0.83, $P < 2.2e{-}16$

CD45+ cell / 100μm²

H&E−based lymphocyte cell / 100μm²

a

Immune cold

Immune hot

b

n=275 region; 85 patients

LUAD   LUSC   Other

Lymphocyte%

CRUK0034 CRUK0037 CRUK0022 CRUK0007 CRUK0025 CRUK0050 CRUK0023 CRUK0051 CRUK0041 CRUK0003 CRUK0036 CRUK0052 CRUK0048 CRUK0001 CRUK0010 CRUK0042 CRUK0029 CRUK0006 CRUK0015 CRUK0005 CRUK0018 CRUK0061 CRUK0035 CRUK0031 CRUK0002 CRUK0044 CRUK0047 CRUK0046 CRUK0055 CRUK0009 CRUK0038 CRUK0024 CRUK0004 CRUK0013 CRUK0053 CRUK0032 CRUK0016 CRUK0021 CRUK0028 CRUK0008 CRUK0020 CRUK0017 CRUK0012 CRUK0057 CRUK0045 CRUK0026 CRUK0027 CRUK0030 CRUK0087 CRUK0071 CRUK0063 CRUK0077 CRUK0089 CRUK0070 CRUK0092 CRUK0062 CRUK0067 CRUK0086 CRUK0076 CRUK0064 CRUK0079 CRUK0081 CRUK0083 CRUK0075 CRUK0091 CRUK0074 CRUK0085 CRUK0073 CRUK0066 CRUK0069 CRUK0078 CRUK0082 CRUK0088 CRUK0072 CRUK0068 CRUK0100 CRUK0094 CRUK0096 CRUK0099 CRUK0098 CRUK0095

CD8⁺%

CD4⁺FOXP3⁻%

CD4⁺FOXP3⁺%

c

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⁺%   CD4⁺FOXP3⁻%   CD4⁺FOXP3⁺%

d

Estimated infiltration

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

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

a

Genomic distance

0.00036

cold-cold | hot-hot
Immune phenotype

b

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

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

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}

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**

		Hazard ratio	p
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

Hazard ratio

g   **LATTICe-A**

		Hazard ratio	p
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

Hazard ratio

a

b

Fractal dimension

0.00036

cold   hot
Immune phenotype

c

Fractal dimension

0.0057

No loss   Loss
LOH: any HLA type

d

H&E   H&E-based spatial modelling of lymphocytes   IHC-derived immune subsets

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

- ● CD8+
- ● CD4+FOXP3-
- ● CD4+FOXP3+

e

Adjacent–tumor lymphocytes/stroma

0.0074

>=Median   <Median
Clonal neoantigens

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

Y-axis: Adjacent-tumor lymphocytes/stroma
X-axis: 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

Y-axis: Tumor%
X-axis: ASCAT Purity

rho = 0.32, *P* = 0.00031
rho = 0.35, *P* = 0.00075
rho = 0.36, *P* = 7.7e-09

Y-axis: Tumor%
X-axis: 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

Y-axis: Lymphocyte%
X-axis: 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

Y-axis: IHC CD8+%
X-axis: H&E Lymphocyte%

f

IHC
H&E
Serial staining

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

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

Y-axis: Lymphocyte%
X-axis: Lymphocyte cell fraction (scored by pathologist)

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

Y-axis: Adjacent-tumor lymphocytes/stroma
X-axis: Pathology TIL estimates

g

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

Y-axis: SMA+ cell / 100μm²
X-axis: H&E-based stromal cell / 100μm²

Legend: LUAD, LUSC, Other, — All

a

Pathology TIL estimates vs Immune classification (Cold, Hot), p=4.6e−08

b

Deep learning

		Hot	Cold
RNA-seq	High	35	18
	Low	13	43

p=7.8e−06

c

Fraction of immune hotspot vs Classification agreement (Discrepant, In agreement), 0.011

d

Lymphocyte% — cold vs hot, ****
Intra-tumor lym/tumor — cold vs hot, ****
Adjacent-tumor lym/stroma — cold vs hot, ****
ASCAT.purity — cold vs hot, ****

H&E-deep learning region specific immune classification

Lymphocyte% — low vs high, ****
Intra-tumor lym/tumor — low vs high, ****
Adjacent-tumor lym/stroma — low vs high, ***
ASCAT.purity — low vs high, ****

RNA-seq regions specific immune classification