Precision Clinical Medicine, 2025, **8**: pbaf001

DOI: 10.1093/pcmedi/pbaf001

Correspondence

Spatially defined intratumoral immune biomarkers predict recurrent versus second primary tumors in non-small cell lung cancer

Rekha Mudappathi,[1,2,‡] Alanna Maguire,[3,4,‡] Eunhee S. Yi,[5] Yanmei Peng,[1,6] Jennifer M. Kachergus,[7] Andras Khoor,[8] Kexin Tan,[1,9] Isabella Zaniletti,[10,11] Jason A. Wampfler,[12] Yanyan Lou,[13] Pedro A. Reck dos Santos,[14] Jonathan D'Cunha,[15] Zhifu Sun,[12] Li Liu [2], Diane F. Jelinek,[16] Junwen Wang [1,17] Henry D. Tazelaar,[4] E.A. Thompson,[7] Ping Yang [1,*]

[1]Division of Epidemiology, Department of Quantitative Health Sciences, Mayo Clinic, Scottsdale, AZ 85259, USA
[2]College of Health Solutions, Arizona State University, Phoenix , AZ 85004, USA
[3]Department of Cancer Biology, Mayo Clinic, Scottsdale, AZ 85259, USA
[4]Department of Laboratory Medicine and Pathology, Mayo Clinic, Scottsdale, AZ 85259, USA
[5]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN 55905, USA
[6]Department of Oncology, Fangshan Hospital, Beijing University of Chinese Medicine, Beijing 102400, China
[7]Department of Cancer Biology, Mayo Clinic, Jacksonville, FL 32224, USA
[8]Department of Laboratory Medicine and Pathology, Mayo Clinic, Jacksonville, FL 32224, USA
[9]Department of integrated Chinese and Western medicine, the Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University/Hunan Cancer Hospital, Changsha 410013, China
[10]IZ Statistics LLC , Tampa, FL 33647, USA
[11]Division of Biostatistics and Computational Biology, Department of Quantitative Health Sciences, Mayo Clinic, Scottsdale, AZ 85259, USA
[12]Division of Clinical Trials and Biostatistics, Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN 55905, USA
[13]Department of Oncology, Mayo Clinic, Jacksonville, FL 32224, USA
[14]Department of Cardiothoracic Surgery, Mayo Clinic, Phoenix, AZ 85054, USA
[15]Department of Hematology/Oncology, Mayo Clinic, Jacksonville, FL 32224, USA
[16]Department of Immunology, Mayo Clinic, Scottsdale, AZ 85259, USA
[17]Division of Applied Oral Sciences & Community Dental Care, Faculty of Dentistry, The University of Hong Kong, Hong Kong 999077, China
*Corresponding author: Ping Yang, yang.ping@mayo.edu
‡Rekha Mudappathi and Alanna Maguire contributed equally to this work.

Dear Editor,

In patients with early-stage non-small cell lung cancer (NSCLC), the risk of post-surgical recurrence (REC) is 20%–50%, [1] and the risk of developing a second primary lung cancer (2P) is 3–4 times higher than in the general population [2]. Distinguishing between REC and 2P can be challenging, complicating treatment of these so called "equivocal" (EQU) cases. We hypothesize that the tumor immune profile can provide insight into the immune system's role in controlling micro-metastasis and suppressing recurrent and/or new primary tumors. In this pilot study, we employed the NanoString GeoMx® Digital Spatial Profiler (GeoMxDSP) to detect intratumoral RNA and protein abundance by region of interest (ROI) [3]. Exemplary H&E-stained tumor tissues were pathologically reviewed; two types of ROI were identified here (Fig. 1A): inflamed (hot) defined as microscopically infiltrated with ≥5% of lymphocytes, macrophages, and plasma cells collectively [4], or uninflamed (cold). All ROIs were further segmented by immunophenotyping (PanCK+, CD45+), where PanCK+ and CD45+ cells respectively represented epithelial tumor and tumor immune microenvironment (TIME) cells. We herein report how to identify intratumoral immune response RNA and protein biomarkers (immune targets) that predict REC and 2P NSCLC.

A total of 62 patients were identified from an ongoing patient cohort [5] for this study if they had a surgically resected, primary adenocarcinoma or squamous cell carcinoma of the lung, pathologic staged T1-2N0M0, with or without a subsequent event of either a recurrence [6] or a second primary tumor [7]. These patients had a mean diagnosis age of 61 (± standard deviation 13) years, were 52% female, 42% never-smokers, 76% non-Hispanic White, 84% adenocarcinoma and 16% squamous cell carcinoma (supplementary Tables S1 and S2, see online supplementary material). Two independent groups served as the discovery [Set-1 ($n = 38$)] and validation [Set-2 ($n = 24$)] sets. The decision of REC ($n = 6$) and 2P ($n = 6$) tumors were abstracted from medical records with a multidisciplinary-consensus diagnosis, based on conclusive reports from pathology, radiology, thoracic surgery, medical oncology, and pulmonary medicine. The five EQU cases are those that cannot be unequivocally categorized as a REC or 2P.

Overall, 847 segments (421 PanCK+, 426 CD45+) among 362 ROIs derived from 62 NSCLC tumor tissues were analyzed. Two serial FFPE tissues were analyzed using GeoMxDSP RNA and protein nCounter panels respectively (supplementary Tables 3a and 3b, see online supplementary material), including the Immune Pathways Panel Human RNA core (73 genes, 5 housekeepers, 6 control probes), and 7 NanoString GeoMx protein modules targeting a total of 71 proteins, including 3 housekeeping controls and

Received 14 November 2024; revised 31 December 2024; accepted 15 January 2025. published 17 January 2025

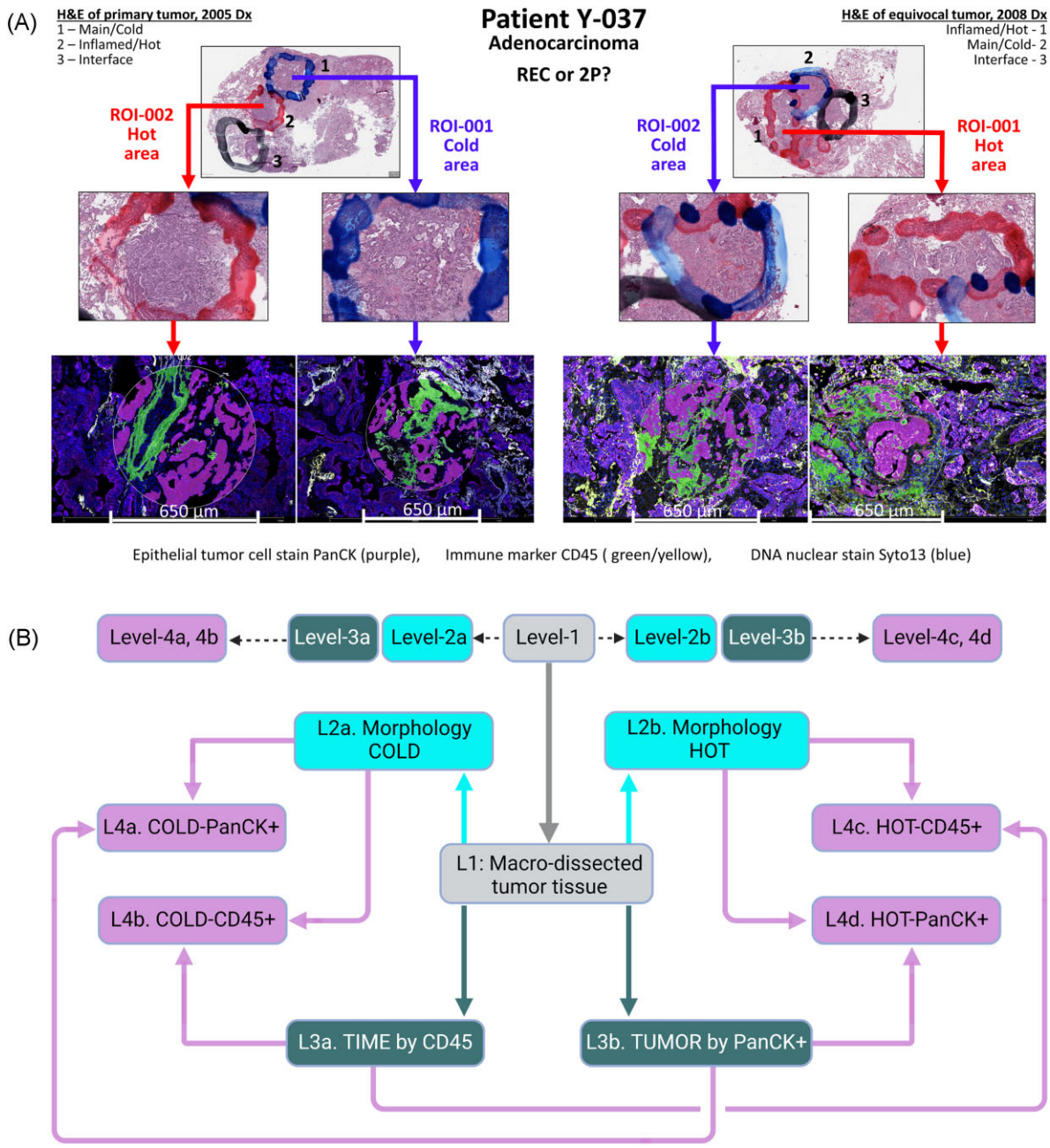

Figure 1. Identification of intratumoral immune response biomarkers that predict REC and 2P NSCLC. (**A**) Is the 2nd tumor (right) a recurrence or a new primary? This patient developed two lung adenocarcinomas 3 years apart; H&E-stained tumor sections are shown in the top panels. Middle panels: color-marked Inflamed/hot (red) or main/cold (blue) areas within each tumor section, mimicking macro-dissected tissue samples, with differing immune cell infiltrations. Bottom panels show two spatial segments or compartments within each ROI: PanCK + TUMOR and CD45 + TIME. (**B**) Analyses and model builds. The four spatially defined tissue sector analyses are outlined in 4-levels. Level-1: microscopically inspected tumor tissue areas (hot and cold) including both ROI subgroups PanCK+ (TUMOR) and CD45+ (TIME) cells, akin to macro-dissected bulk tumor samples (grey). Level-2: morphologically distinct tumor tissue areas, either hot or cold including both PanCk+ (TUMOR) and CD45+ (TIME) cells (cyan). Level-3: segment-distinct tumor tissue areas either TUMOR or TIME cells (dark green). Level-4: morphology- and segment-specific tumor tissue areas: cold-TUMOR, cold-TIME, hot-TUMOR, hot-TIME (pink). (**C**) Validated Immune functional markers predicting outcomes of stage-I NSCLC. The two hemispheres, color shaded to represent specific spatial designations, show significant markers for REC (left) or 2P (right). Protective and hazardous markers are shown in blue and red text respectively. (**D**) Predicting alignment of EQU to REC or 2P using 13 validated protein markers. All four EQU cases (Y-035, Y-036, Y-037, and Y-038) had subsequent tumors that arose over 2 years (4.5, 2.1, 3.5, and 2.1 years respectively) later and with the same histology, but in a different lung lobe to the primary tumors assessed in this study. The same histology supports recurrence, while the different lung lobe supports the subsequent tumor being a second primary tumor. Typically, the timeframe for REC and 2P is considered to arise within or after 2 years.

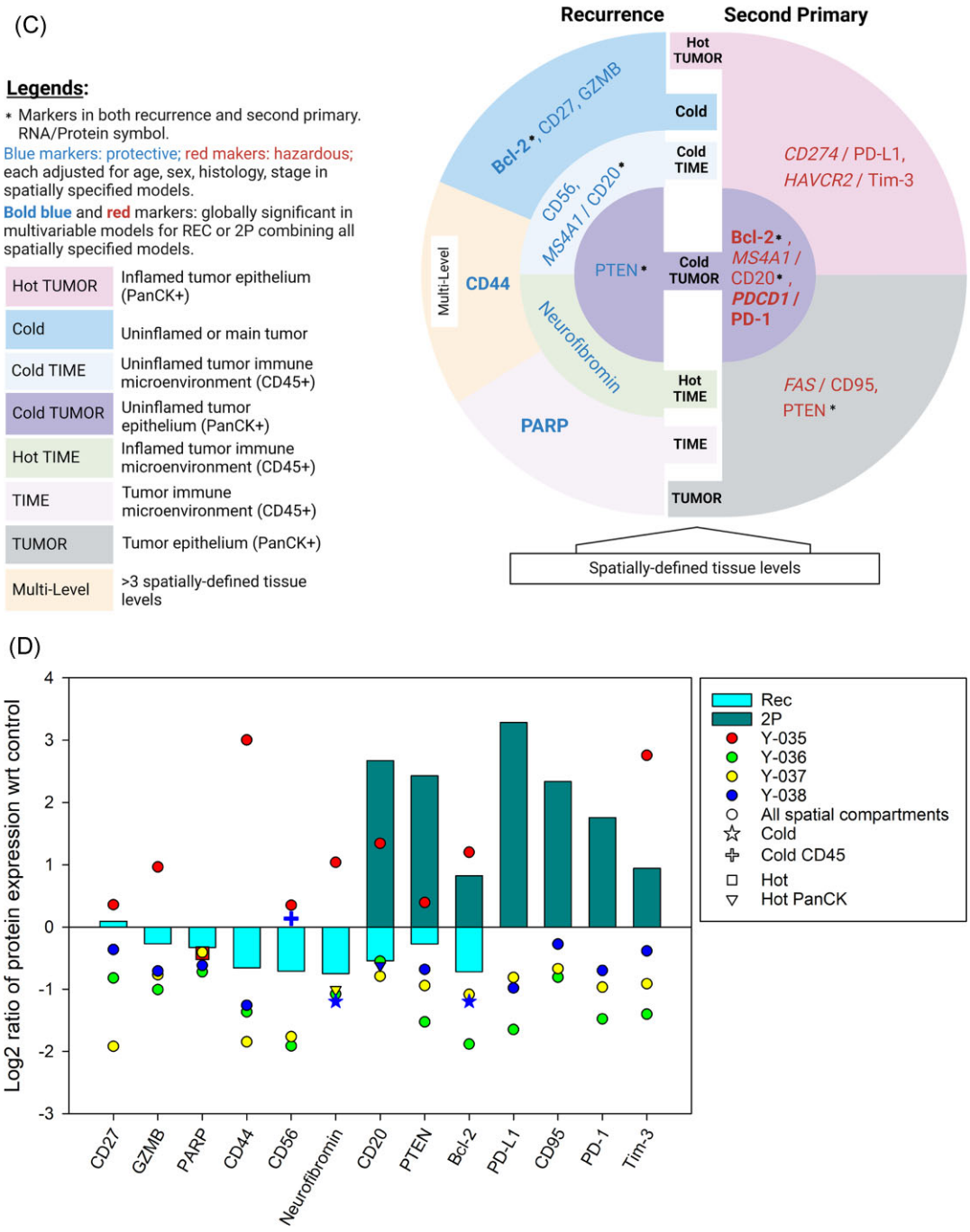

Figure 1. *Continued.*

3 negative controls. Cell counts as per SYTO-13-positive nuclei counts are provided in supplementary Table S4 (see online supplementary material). Set-2 only had protein data because the RNA platform was not available. The overall distributions of the log2-transformed expression level of all RNA and protein targets are illustrated in supplementary Fig. S1a–c, see online supplementary material.

Four-level bi-directional discovery-validation analyses (Fig. 1B) were conducted between RNA and protein, and two independent protein datasets. Cox proportional hazard (Cox) models for predicting REC and 2P were built at four spatial levels starting with univariate model of all ROIs for all RNA markers and eight covariates (age, sex, smoking history, race, histology, stage, morphology, segment). The 2P-smk model in lieu of 2P was used because zero "never-smokers" developed 2P. Cox models were clustered by unique patient ID to account for intra-person correlation due to multiple ROIs per tumor sample [9]. Hazard ratio (HR) values with 95% confidence interval (CI) were estimated for each biomarker and covariate, with those having $P < 0.2$ considered for inclusion in the multivariable Cox model. Multiple testing corrections [10] utilized false discovery rate (FDR) <0.01 threshold in all discovery models (to minimize overlapping markers between REC and 2P) and FDR < 0.2 in validation models. In addition to frequency-matched confounding effects by selected known prognostic factors in both sample sets, residual confounding effects were tested at all four levels of ROI-based analyses and for both RNAs and

protein markers. Potential confounders [Supplementary Fig. S2 (see online supplementary materials), covariates in green box] were adjusted. A 2-tailed *P*-value < 0.05 was considered statistically significant in the final (global) multiple Cox models, adjusting for significant covariates identified at various spatial levels. All statistical analyses were conducted using the R software (v4.1.2).

In Level-1, all cold and hot ROIs with both TUMOR and TIME cells, resembling bulk macro-dissected tumor tissue, were included. Six independent multivariable Cox models were built (3 for REC, 3 for 2P) following the bi-directional discovery-validation process (supplementary Table S5, see online supplementary material). Then, ROIs that were spatially defined by morphology (hot, cold), segment (PanCK+, CD45+), and morphology plus segment (hot PanCK+, hot CD45+, cold PanCK+, cold CD45+) were assessed, yielding 24 multivariable Cox models: 12 for REC and 12 for 2P. Of these, 6 were built at the morphology level (hot vs. cold), 6 at the segment level (PanCK+ vs. CD45+), and 12 at the morphology by segment level (supplementary Tables S6 and S7, see online supplementary material). Each validated marker was interrogated against the published literature to determine its novelty in predicting REC and/or 2P. A total of 130 publications and 4324 candidate markers were interrogated; of these, 4 articles reported markers validated in our study, but none by GeoMxDSP technology (supplementary Table S8, see online supplementary material). We identified predictive markers for REC and 2P at different spatially defined tumor areas and cell designations (Fig. 1C). The two hemispheres show significant markers for REC (left) or 2P (right). Protective markers are shown in blue text, hazardous in red. Markers that predicted both REC and 2P are intriguing: Bcl-2 and CD20 were protective for REC when detected at cold or cold-CD45 + levels but hazardous for 2P when detected at the cold-PanCK+ level, showing tumor-spatial-marker specificity

Next, we developed two global predictive models acoss all four levels, one for REC and one 2P; significant markers are in bold font (Fig. 1C), revealing nine novel prognostic biomarkers: three that protect against REC, four hazardous for 2P, and two that exhibit opposing effects between REC and 2P. Notably, reduced expression of CD44 and PARP were found to be protective against REC, while increased RNA and protein expression of *PDCD1*/PD-1 were found to be hazardous for 2P. Interestingly, reduced protein expression of BCL2 was found to be protective for REC, while increased protein expression was hazardous for 2P.

Finally, EQU cases were compared to the validated protein markers from Set-1 with the objective of predicting REC or 2P classifications. The normalized log2-transformed expression ratios for each protein target within each group (none, REC, 2P) and EQU case compared to the control group (none) was determined. Bioinformatically validated targets for which an EQU case had expression equal, exceeding (negative or positive), or contradictory to the mean REC or 2P group expression were visualized using SigmaPlot 14.5. The log2-transformed REC/none and 2P/none ratios are depicted as histograms and represent the mean expression for protein targets identified as significant within REC or 2P, Fig. 1D. Using our prototypic "Alignment Score" algorithm (supplementary Table S9, see online supplementary material) as a proof-of-principle, by either a simple count of supporting markers (equal or exceeding the reference mean) or penalized points (subtracting contradictory markers from the supporting), allowed us to align the EQU cases. The results show that Y-035 (red symbol) aligns to 2P, while Y-037 (yellow), Y-036 (green) and Y-038 (blue) align with REC. In summary, our findings demonstrated the feasibility of identifying differential biomarkers that predict REC or 2P at spatially defined intra-tumoral sectors, supporting new directions in immunotherapies and interceptions beyond the existing strategies.

Acknowledgments

This study was supported by the National Cancer Institute (Grants No. R01-80127, Yang and R01-84354, P.Y.) and the National Institutes of Health (Grant No. LM-R01-13438, L.L., J.W., and P.Y.). We thank J. Shi from the Department of Cancer Biology of Mayo Clinic.

Author contributions

Rekha Mudappathi (Formal analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing), Alanna Maguire (Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing), Eunhee S. Yi (Data curation, Investigation, Writing – review & editing), Yanmei Peng (Data curation, Investigation, Writing – review & editing), Jennifer M. Kachergus (Data curation, Investigation, Visualization, Writing – review & editing), Andras Khoor (Data curation, Writing – review & editing), Kexin Tan (Data curation, Investigation, Writing – review & editing), Isabella Zaniletti (Formal analysis, Investigation, Visualization, Writing – review & editing), Jason A. Wampfler (Data curation, Formal analysis, Investigation, Methodology, Writing – review & editing), Yanyan Lou (Data curation, Writing – review & editing), Pedro A. Reck dos Santos (Conceptualization, Writing – review & editing), Jonathan D'Cunha (Conceptualization, Writing – review & editing), Zhifu Sun (Conceptualization, Writing – review & editing), Li Liu (Conceptualization, Writing – review & editing), Diane F. Jelinek (Conceptualization, Writing – review & editing), Junwen Wang (Conceptualization, Methodology, Writing – review & editing), Henry D. Tazelaar (Conceptualization, Data curation, Writing – review & editing), E.A. Thompson (Conceptualization, Data curation, Writing – review & editing), and Ping Yang (Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing).

Supplementary data

Supplementary data is available at *PCMEDI Journal* online.

Conflict of interest

None declared. In addition, as an Editorial Board Member of *Precision Clinical Medicine*, the corresponding author Ping Yang was blinded from reviewing and making decision on this manuscript.

References

1. Potter AL, Costantino CL, Suliman RA *et al.* Recurrence after complete resection for non-small cell lung cancer in the national lung screening trial. *Ann Thorac Surg* 2023;**116**:684–92. https://doi.org/10.1016/j.athoracsur.2023.06.004.
2. Choi E, Luo SJ, Ding VY *et al.* Risk model-based management for second primary lung cancer among lung cancer survivors through a validated risk prediction model. *Cancer* 2024;**130**:770–80. https://doi.org/10.1002/cncr.35069.
3. Steele LE, Tan CW, Chew KY *et al.* Unlocking spatially-resolved transcriptomic and proteomic secrets of century-old lungs from the 1918 'Spanish' Influenza Pandemic. NanoString Website.

2024. Available at: https://nanostring.com/wp-content/uploads/2024/02/AGBT-2024_PST-116-Kirsty-Short_U-Queensland_AGBT-1918-lung-poster_v2.pdf (31 December, 2024, date last accessed).
4. Wang L, Geng H, Liu Y *et al.* Hot and cold tumors: immunological features and the therapeutic strategies. *MedComm* 2023;**4**:e343. https://doi.org/10.1002/mco2.343.
5. Xie D, Allen MS, Marks R *et al.* Nomogram prediction of overall survival for patients with non-small-cell lung cancer incorporating pretreatment peripheral blood markers. *Eur J Cardiothorac Surg* 2018;**53**:1214–22. https://doi.org/10.1093/ejcts/ezx462.
6. Antakli T, Schaefer RF, Rutherford JE *et al.* Second primary lung cancer. *Ann Thorac Surg* 1995;**59**:863–6. discussion 867. https://doi.org/10.1016/0003-4975(95)00067-U.
7. Mahul B, Amin DMG, Meyer Vega LR *et al.* Compton. *American Joint Committee Cancer Staging Manual*, 8th edition. Springer 2017.
8. Merritt CR, Ong GT, Church SE *et al.* Multiplex digital spatial profiling of proteins and RNA in fixed tissue. *Nat Biotechnol* 2020;**38**:586–99. https://doi.org/10.1038/s41587-020-0472-9.
9. Chen G, Taylor PA, Haller SP *et al.* Intraclass correlation: improved modeling approaches and applications for neuroimaging. *Hum Brain Mapp* 2018;**39**:1187–206. https://doi.org/10.1002/hbm.23909.
10. Benjamini Y, Hochberg Y. . Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* 1995;**57**:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Received 14 November 2024; revised 31 December 2024; accepted 15 January 2025. published 17 January 2025