



Unveiling the influence of tumor and immune signatures on immune checkpoint therapy in advanced lung cancer

Nayoung Kim, Sehhoon Park, Areum Jo, Hye Hyeon Eum, Hong Kwan Kim, Kyungjong Lee, Jong Ho Cho, Bo Mi Ku, Hyun Ae Jung, Jong-Mu Sun, Se-Hoon Lee, Jin Seok Ahn, Jung-Il Lee, Jung Won Choi, Dasom Jeong, Minsu Na, Huiram Kang, Jeong Yeon Kim, Jung Kyoony Choi, Hae-Ock Lee , Myung-Ju Ahn 

Department of Microbiology, College of Medicine, The Catholic University of Korea, Seoul, 06591, Korea • Department of Biomedicine and Health Sciences, Graduate School, The Catholic University of Korea, Seoul, 06591, Korea • Division of Haematology-Oncology, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, 06351, Korea • Department of Thoracic and Cardiovascular Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, 06351, Korea • Division of Pulmonary and Critical Care Medicine, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, 06351, Seoul, Korea • Research Institute for Future Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, 06351, Korea • Department of Neurosurgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, 06351, Republic of Korea • Department of Bio and Brain Engineering, KAIST, Daejeon, Korea • Precision Medicine Research Center, College of Medicine, The Catholic University of Korea, Seoul, 06591, Korea

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

Abstract

This study investigates the variability among patients with non-small cell lung cancer (NSCLC) in their responses to immune checkpoint inhibitors (ICI). Recognizing that patients with advanced-stage NSCLC rarely qualify for surgical interventions, it becomes crucial to identify biomarkers that influence responses to ICI therapy. We conducted an analysis of single-cell transcriptomes from 33 lung cancer biopsy samples, with a particular focus on 14 core samples taken before the initiation of palliative ICI treatment. Our objective was to link tumor and immune cell profiles with patient responses to ICI. We discovered that ICI non-responders exhibited a higher presence of CD4⁺ regulatory T cells, resident memory T cells, and TH17 cells. This contrasts with the diverse activated CD8⁺ T cells found in responders. Furthermore, tumor cells in non-responders frequently showed heightened transcriptional activity in the NF- κ B and STAT3 pathways, suggesting a potential inherent resistance to ICI therapy. Through the integration of immune cell profiles and tumor molecular signatures, we achieved an accuracy rate exceeding 95% in predicting patient responses to ICI treatment. These results underscore the crucial importance of the interplay between tumor and immune microenvironment, including within metastatic sites, in affecting the effectiveness of ICIs in NSCLC.

eLife assessment

The authors utilized scRNAseq profiling of NSCLC patient tumor samples to generate **useful** insights into the determinants of ICI responsiveness in NSCLC patients. While some of the findings add weight to the current literature, the analysis is **incomplete** due to the small cohort size and occasional departures from recognized subtype markers. This study would benefit from external cohorts to both validate the findings and to justify the statistical analysis undertaken.

<https://doi.org/10.7554/eLife.98366.1.sa2>

Introduction

Treatment landscape in cancer has rapidly evolved with the introduction of immune checkpoint inhibitors (ICI). Among various immune checkpoints, antibody-targeting programmed cell death-1 (PD-1) and its ligand (PD-L1) have demonstrated clinical benefits over conventional systemic chemotherapy in patients with non-small cell lung cancer (NSCLC). They have been approved as either monotherapy in patients with high PD-L1 expression¹ or combined with cytotoxic chemotherapy regardless of PD-L1 expression^{2,3}. Moreover, the clinical benefits were validated in unresectable stage III NSCLC as consolidation therapy after definitive chemoradiotherapy or early stage NSCLC (Ib-IIIA) as adjuvant therapy after curative surgery⁴.

There have been many efforts to elucidate the predictive biomarkers of PD-(L)1 inhibitors. The PD-L1 expression in tumor tissue evaluated via immunohistochemistry has been incorporated as a companion diagnostic biomarker from the early clinical trials, which enhanced the response rate up to 46% in patients with PD-L1 $\geq 50\%$ and showed an overall survival rate of up to 26.3 months⁵. Other biomarkers, such as tumor mutation burden or gene expression profile, also demonstrated a positive predictive value^{6,7}. Recent large-scale meta-analysis of clinico-immunogenomics shows that both tumor- and T cell intrinsic factors exert a substantial impact on ICI response⁸, supporting the necessity of in-depth investigation of high-throughput profiles of tumor and microenvironment.

ICI treatment modifies systemic immune profiles represented by changing the proportion of specific cell types. The increase in PD-1+Ki67+CD8+ T cells in the peripheral blood after treatment with PD-1 inhibitors is associated with a better outcome in patients with NSCLC^{9,10}. The cell type has also been identified in tumor tissues and related to clinical outcomes¹¹.

Tumor-infiltrating PD-1 positive T cells have higher capacity of tumor recognition than PD-1 negative T cells, indicating certain types of T cells in tumor bed function as tumor-specific T cells. In addition to T cells, other immune cell types, such as myeloid-derived suppressor cells or regulatory T cells, regulating the tumor-specific T cell immunity may also influence the therapeutic outcome¹²⁻¹⁴. In summary, the multicellular regulation of the tumor-immune microenvironment highlights the importance of systemic tumor and immune cell profiling at single cell resolution to investigate baselines associated with response to ICI treatment.

Results

Variability and features of the lung cancer samples

We conducted scRNA-seq on 33 lung cancer samples from 26 patients treated with immune checkpoint inhibitors (ICI) between August 2017 and December 2019 to understand how cellular dynamics in lung cancer affect treatment sensitivity to PD-(L)1 inhibitors, used alone or in combination (**Fig. 1a** and Table S1). Immune checkpoint therapy provides clinical benefit in advanced metastatic NSCLC across different treatment lines¹⁵. Notably, our samples have been collected from various tissue sites. In the scRNA-seq analysis, all specimens were used for the cell type profiling in an unbiased manner. For the evaluation of clinical outcomes, refined 14 core samples from 11 patients were used to minimize sample specific variations. Exclusion criteria from the core group encompass samples with treatment applied as adjuvant therapy, acquired after ICI treatment, no tumor content, non-evaluable for the clinical response, or histology other than non-small cell lung cancer (**Table 1** and **Fig. 1a**). Of the 11 core patients, 8 had adenocarcinoma and 3 had squamous cell carcinoma. Clinical outcomes of ICI were partial response (PR) in four patients, stable disease (SD) in two patients, and progressive disease (PD) in five patients. Patients were classified as responders (PR) and non-responders (SD and PD) according to ICI response.

Due to the diversity of sample collection sites, our data may be influenced by varied immune cell composition at different sample collection sites. Therefore, we analyzed immune and stromal cell subsets across early-stage (tLung) and late-stage (tL/B) lung tumors, and metastatic lymph nodes (mLN), comparing them to normal lung (nLung) and lymph node (nLN) tissues. This analysis was conducted on public scRNA-seq data from 43 samples from 33 LUAD patients¹⁶ (**Fig. S1a-c**). Although there were differences in tissue-specific resident populations, we found that the immune cell profiles, especially T/NK cells of mLN were similar to those of primary tumor tissues indicating the activation of immune responses were consistently observed at metastatic sites (**Fig. S1d-f**).

Classification of immune cell subset in lung cancer

Global cell-type profiling (**Fig. 1b,c** and Table S2) illustrates the cellular composition of each sample as epithelial/tumor cells, fibroblasts, endothelial cells, T/natural killer (NK) cells, B/plasma cells, myeloid immune cells, and mast cells. Individual samples show variations in epithelial/tumor content as well as in immune cell composition (**Fig. 1d,e**).

For further analysis of immune cell subtypes, we applied sequential subclustering on global immune cell clusters. As scRNA-seq shows limited performance in separating CD4⁺ and CD8⁺ T cell subsets, antibody-derived tag (ADT) information¹⁷ was used to complement the transcriptome data and to predict CD4⁺ T cells, CD8⁺ T cells, and NK cells (**Fig. 2a**). Finally, fourteen CD4⁺ and fourteen CD8⁺ T cell subclusters were identified excluding <5% ambiguous cells (**Fig. 2b,c** and Table S2). In the CD4⁺ T cell compartment, naïve-like T cells (TN, CD4_cluster0) and central memory T cells (TCM, CD4_cluster1) expressing *SELL*, *TCF7*, *LEF1*, and *CCR7* genes or tissue-resident memory T cells (TRM, CD4_clusters3, 5, 6) expressing *NR4A1*, *MYADM*, and *PTGER4* genes were abundant in most samples. Regulatory T cells (Treg, CD4_cluster2) with *FOXP3*, *CTLA4*, *ICOS*, and *BATF* expression were also abundant, which has been demonstrated as tumor-specific alterations in the tissue microenvironment^{16,18,19}. In the CD8⁺ T cell compartment, effector memory T cells (TEM), effector T cells (TEFF), effector memory CD45RA positive cells (TEMRA) (CD8_clusters0, 2, 3, 8) expressing *PRF1* and *IFNG* were dominant over TN/TCM (CD8_clusters4, 5) types. Exhausted T cells (TEX, CD8_clusters1, 12) expressed multiple checkpoint genes (*HAVCR2* and *PDCD1*) along with high levels of *PRF1*, *IFNG*, *CXCR3*, and *CXCL13*. Co-expression of cytotoxic effectors and checkpoint molecules in TEX clusters indicates that cluster populations may retain functional capacity as cytotoxic effector T cells²⁰. Further, clonotype analysis of TCR supported the T cell subset classification demonstrating higher clonal expansion in the CD8⁺ T cell compartment than that in CD4⁺ T cells, with the highest levels within the TEX subclass (**Fig. S2a-c**).

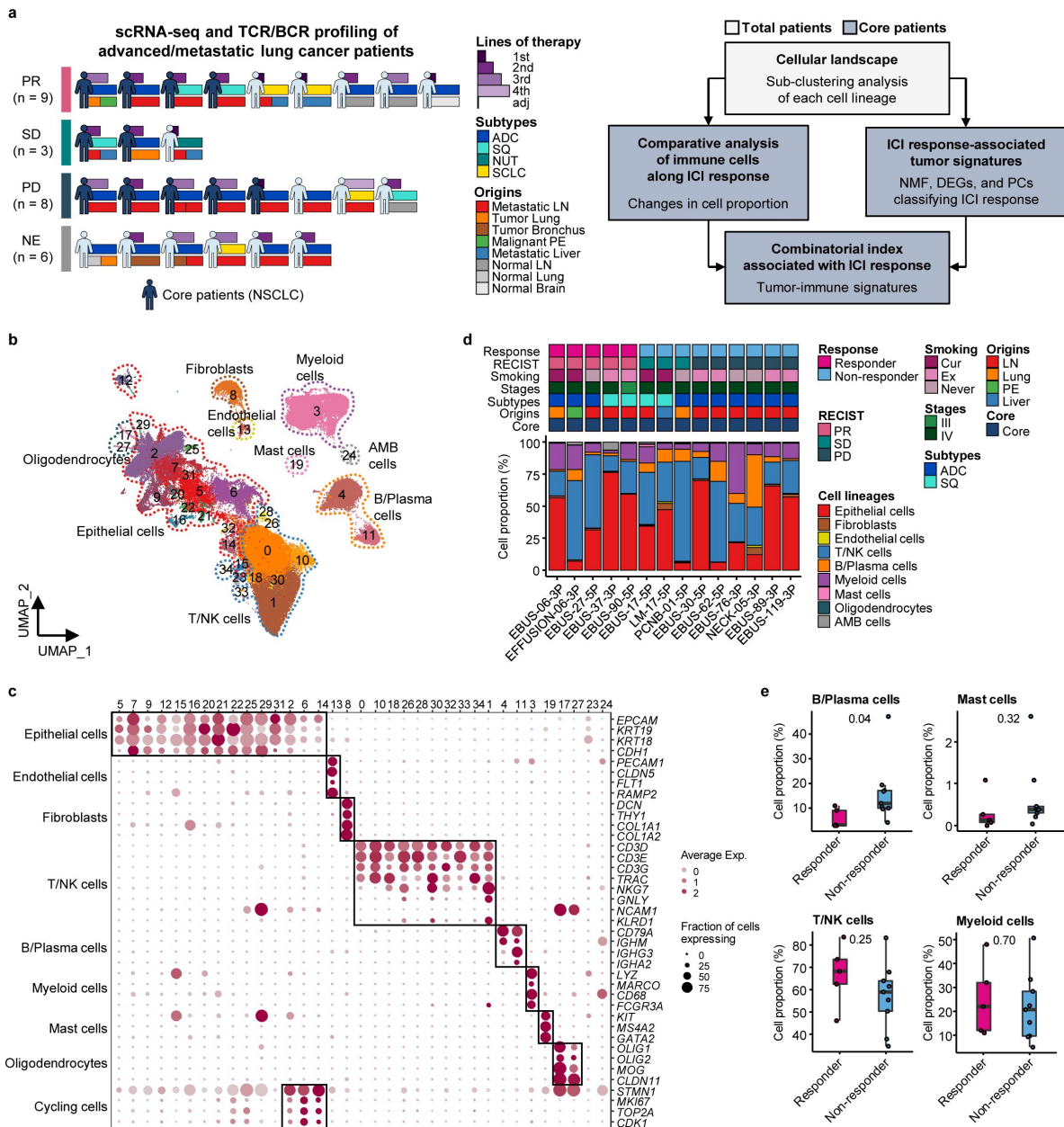


Figure 1.

Cell lineage identification of 96,505 single cells from 26 patients with lung cancer treated with ICI.

a, Workflow of sample collection and single-cell analysis of lung cancer patients treated with ICI. **b**, UMAP plot of 96,505 single cells from 33 samples acquired from 26 advanced lung cancer patients, colored by clusters. AMB cells, Ambiguous cells. **c**, Dot plot of mean expression of canonical marker genes for cell lineages. **d**, Proportions of the cell lineages in NSCLC tissue from core patients shown by individual samples aligned with clinical data. Labels for origins indicate LN, Metastatic lymph node; Lung, Tumor lung; PE, Malignant pleural effusion; Liver, Metastatic liver. **e**, Box plot of the percentage of cell lineages in responder and non-responder groups. Label represents p-value calculated via two-tailed Student's t-test. Each box represents the median and the interquartile range (IQR, the range between the 25th and 75th percentile), whiskers indicate the 1.5 times of IQR.

Immunotherapy	Targets	No. of samples	No. of patients
Pembrolizumab	PD-1	9	8
Atezolizumab	PD-L1	1	1
Nivolumab	PD-1	2	1
Vibostolimab+Pembrolizumab	TIGIT + PD-1	2	1

RECIST	Description	No. of samples	No. of patients
PR	Partial response	5	4
SD	Stable disease	3	2
PD	Progressive disease	6	5

Table 1.

Clinical overview of NSCLC patients treated with ICI.

Natural killer cells can be subclassified as CD56^{bright}, transitional, active, and mature types [21](#) (Fig. S3a,b and Table S2). Active NK cells expressed the highest level of *PRF1*, *TNF*, and *IFNG*, reflecting a cytotoxic effector function.

Compared to T cell clusters, fewer B/plasma cells were detected as follicular (B_clusters 0, 1, 2, 4, 6, 11), germinal center (B_cluster 10), and plasma/mucosa-associated lymphoid tissue (MALT) B cells (B_clusters 3, 7, 12) (Fig. S3a,b and Table S2). Plasma/MALT B cells manifested higher levels of clonal expansion of BCR than follicular B cells (Fig. S3c,d). Identical clonality of some follicular B and plasma cells suggests in situ maturation and differentiation of B cells to plasma cells in tumor tissues (Fig. S3c, clonotype 16).

Myeloid cells were composed of monocytes, dendritic cells, and a large number of macrophages (Fig. S3a,b and Table S2). CD14⁺CD16⁻classical monocytes (Myeloid_clusters1,6) were predominantly found over non-classical CD14^{lo}CD16⁺ types (Myeloid_cluster15). Alveolar macrophages (Alveolar Mac, Myeloid_cluster0) expressed well-defined marker genes such as *MARCO*, *FABP4*, and *MCEMP1* along with anti-inflammatory genes such as *CD163*, *APOE*, and *C1QA/B/C*. Monocyte-derived macrophages represent heterogeneous populations with a similar gene expression profile to alveolar macrophages (Mo-Mac, Myeloid_clusters2, 3, 5, 9, 11, 13), with an elevated chemokine gene expression (CXCL10⁺ Mo-Mac, Myeloid_clusters4, 14, 16), or with active cell cycle progression (Proliferating Mac, Myeloid_cluster7). Dendritic cells were categorized as CD1c⁺ (Myeloid_cluster8, *CD1C* and *ITGAX*), activated (Myeloid_cluster17, *CCR7* and *LAMP3*), and CD141⁺ (Myeloid_cluster19, *CLEC9A* and *XCR1*) subclusters.

Overall immune cell composition is comparable to those reported in previous studies [16](#), [18](#), [19](#).

Immune cell landscape fostering the ICI response

In global cell-type profiling (Fig. 1b,c), abundance in T, NK, or myeloid cell types shows no difference between responders and non-responders (Fig. 1d, e). Nonetheless, as specific differentiation features within the cell types may influence the response to ICI treatment, we compared the response groups using the proportion of subclusters within CD4⁺ T, CD8⁺ T, NK, B/plasma, and myeloid cells. After subclustering (Fig. 2b and Fig. S3a), three subsets of CD4⁺ T cells, i.e., Treg, TRM, and CD4⁺ T helper 17 (TH17), were significantly ($p < 0.01$) overrepresented in the non-responder group (Fig. 2d). In contrast, among CD8⁺ T cell populations, TEM subsets demonstrated a modest level of association with the patients who responded well against progressive disease (Fig. 2d, $p = 0.06$). TCR clonotype analysis supported the cellular dynamics such that clonal expansion was more prominent in cytotoxic CD8⁺ T cells over CD4⁺ Tregs in the responder group (Fig. 2e and Fig. S2c). Overall landscape in each cell type (Fig. 2f) suggests that CD4⁺ Treg and TRM as well as follicular B cells may interfere with the ICI response, whereas CD8⁺ T cell activation (TEM, TEMRA/TEFF, and TEX), mature NK cells, and CXCL10⁺ Mo-Mac cells support the ICI response. The balance between separate immune cell types informs immune regulatory axes that may be targeted to favor the activation of tumor-reactive immunity.

Systemic evaluation of the immune microenvironment associated with ICI response

Next, we evaluated the immune microenvironment as an entity by using all immune cells as a denominator in the subtype proportions. In this setting, we used diverse clinical group comparisons and identified the immune cell blocks separated by clinical outcomes (Fig. 3a). The immune cell blocks overrepresented in the non-responder groups consisted of CD4⁺ Treg, follicular B cells, and CD4⁺ TH17/TRM/ T helper 1 (TH1)-like cells. In the immune cell blocks of the responder groups, CD8⁺ TEM cells showed the strongest enrichment along with the other CD8⁺

TEX/TEMRA/TEFF/Mitochondria (MT) high cells as well as CXCL10+ Mo-Mac. Despite the immune footprints of the ICI responders, extensive variations among individual patients (**Fig. 3b**) hamper patient stratification solely based on the immune profiles.

Tumor cell signatures associated with ICI response

We investigated the associations between genomic characteristics in tumor and ICI response. The constraints of mutation analysis with 10x chromium data complicated the direct correlation between tumor mutation burden and ICI outcome. Rather, we assessed copy number alterations (CNA) indirectly, via chromosomal gene expression patterns. These analyses revealed a moderate correlation between low levels of CNA, including both gain and loss of heterozygosity, and positive responses to ICI (**Fig. S4**). This finding is consistent with the result from previous genetic association studies²².

To assess gene expression characteristics of tumors influencing ICI response, we separated malignant tumor cell clusters from normal epithelial cell types (**Fig. S5**). Subsequent DEG analysis (**Fig. 4a** and **Table S3**) identified genes in poor response groups linked to the regulation of cell death, cell motility, and cell activation (**Fig. S6** and **Table S4**). The DEGs were refined later by combinations of various tumor signatures separating responder and non-responder groups (**Table S5**).

Next, to explore the existence of gene programs and modules influencing the ICI response, we applied factorization using non-negative matrix factorization (NMF) and scINSIGHT²³. Among 30 factors from NMF across all malignant cells, we identified factors showing high loadings for a specific RECIST group as NMF programs p1~4 (**Fig. 4b, c**). There were clear distinction among RECIST groups according to the gene expression levels associated with these NMF programs (**Fig. 4d** and **Table S5**). To identify gene modules consistent across different patients, we examined RECIST-specific modules through scINSIGHT analysis²³ (**Fig. 4e**). Unfortunately, we found that contributions to these gene modules varied significantly among patients. To mitigate this variability, we adjusted the RECIST-specific modules by combining genes from the original modules (**Table S5**). The refined gene modules showed a specific gene expression pattern for each RECIST group, similar to the NMF programs (**Fig. 4f**). Overall, both genes and their functional categories segregated depending on the selection techniques used (**Fig. 4g, h**). However, transcription factors governing the signatures derived from DEG, NMF, and scINSIGHT analyses consistently delineated between responders and non-responders (**Fig. 4i**). Responder-specific gene signatures showed associations with the transcription factors Regulatory Factor X Associated Ankyrin Containing Protein (RFXANK), Regulatory Factor X Associated Protein (RFXAP), and Regulatory Factor X5 (RFX5). This RFX protein complex has emerged as a positive biomarker for the immune response in diverse cancer types²⁴. Non-responder-specific gene signatures were regulated by Activator Of Transcription 3 (STAT3) and Nuclear Factor Kappa B Subunit 1 (NFKB1), known to play roles in PD-L1 regulation and T cell activation in cancer²⁵.

We also adopted principal component analysis (PCA) to isolate correlated gene signatures variably expressed in tumor cells (**Fig. S7**, and **Table S4, S5**). Among the top 10 PCs, negatively correlated genes in PC2, PC7, and PC8 distinguished the tumor cells in the poor response groups, whereas positively correlated genes in PC6 and PC9 were upregulated in the better response groups (**Fig. S8a-c**). Tumor cells from the PR group had low PC2.neg scores, suggesting low growth factor/type I interferon response signaling as a tumor cell-specific positive predictor of the ICI response. Conversely, high levels of growth factor/type I interferon response signaling in tumor cells may present intrinsic resistance to PD-(L)1 inhibitor alone or in combination. Type I interferon is known to drive anti-tumor effect directly or indirectly on tumor and surrounding immune cells, but also acts to counter the anti-tumor effect by inducing CD8+ T cell exhaustion and up-regulating immune-suppressive genes on tumor cells²⁶.

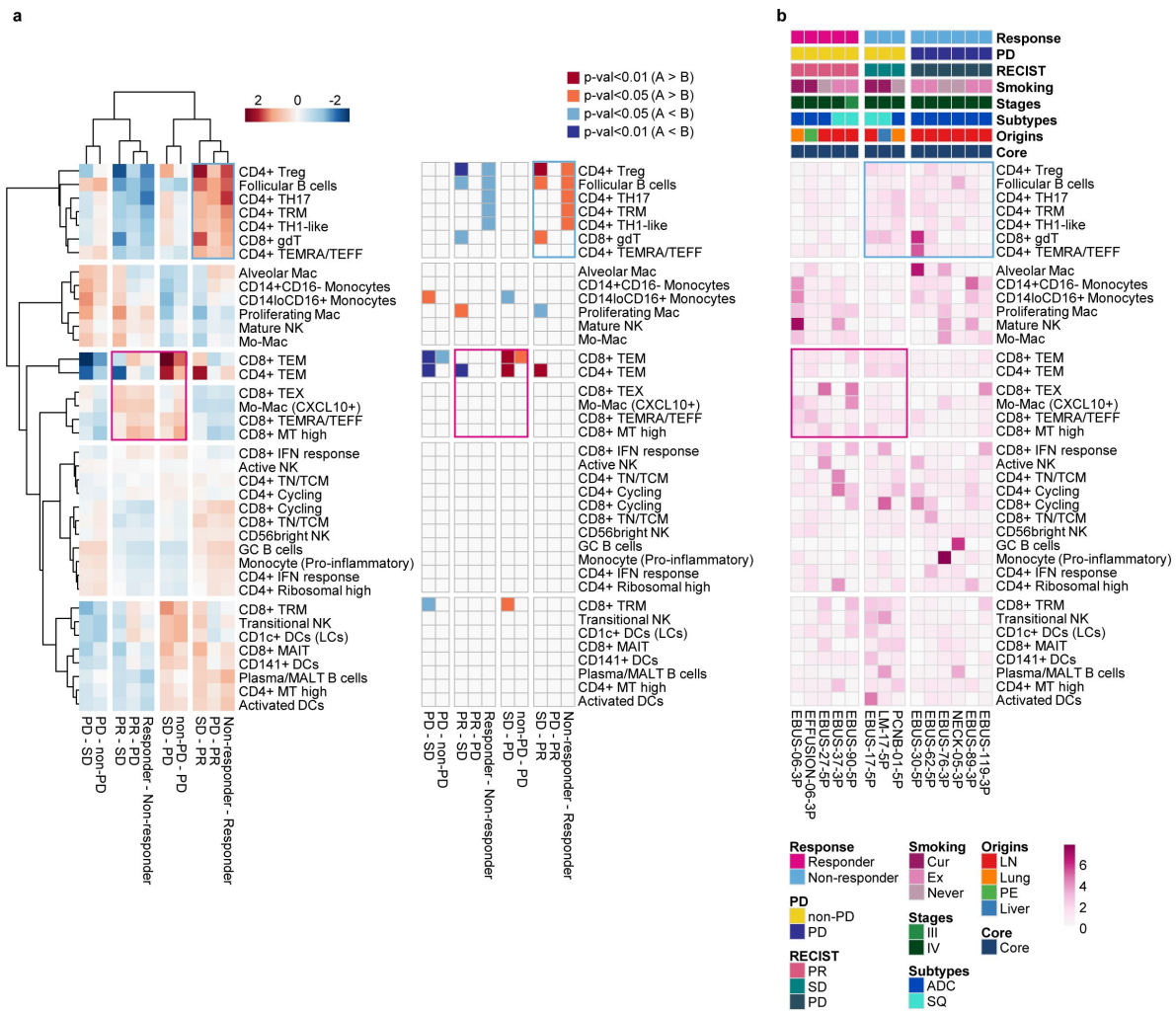


Figure 3.

Systemic evaluation of immune cell dynamics associated with response to ICI.

a, Heat map with unsupervised hierarchical clustering (left) and depicting significance (right) of proportional changes in cell subtypes within total immune cells. Proportional changes were compared for multiple ICI response groups. Color represents the $-\log(p\text{-value})$ determined using two-tailed Student's t-test. **b**, Distribution map for each cell type across individual samples aligned with clinical data. Color represents R_o/e score calculated using the chi-square test.

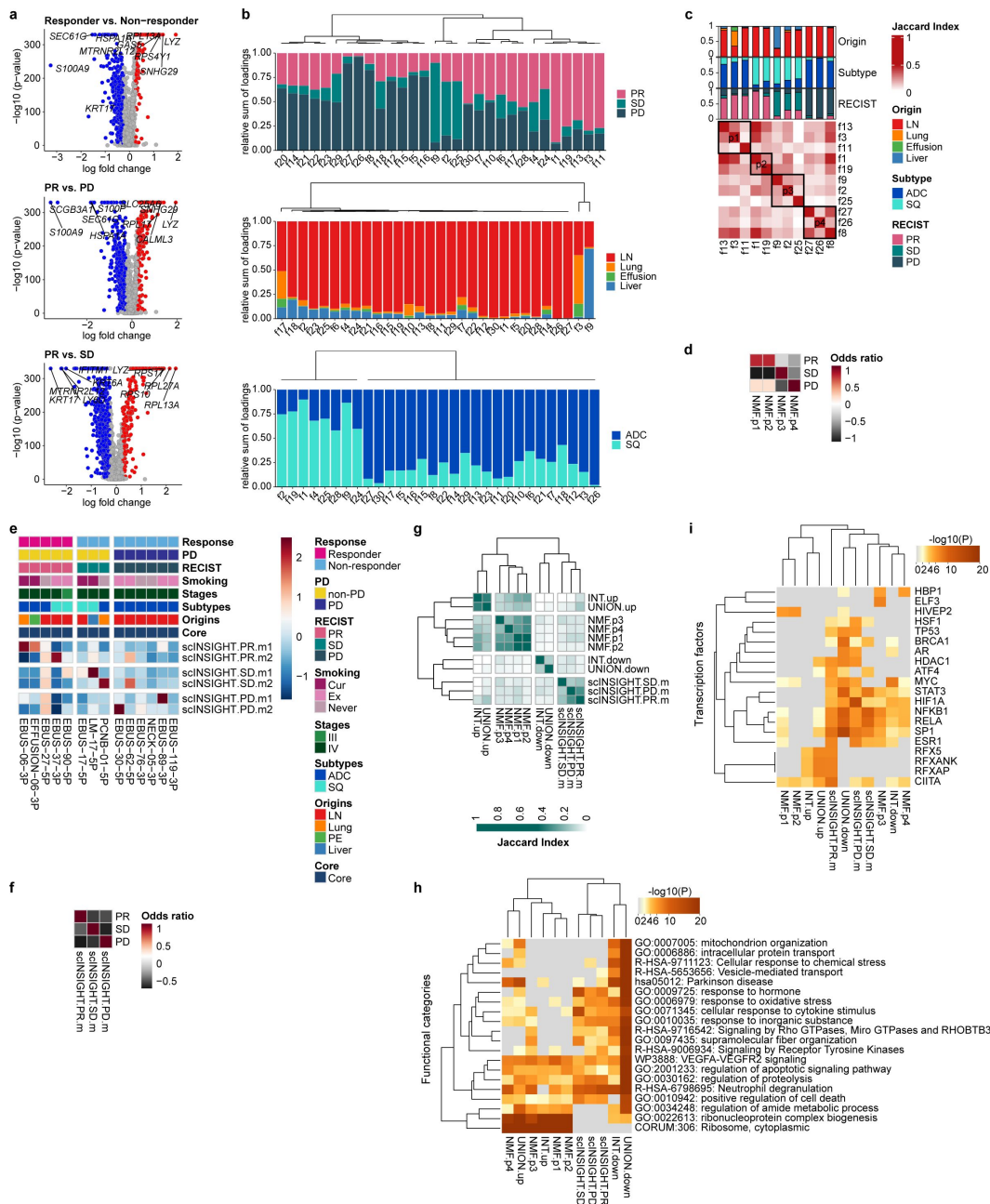


Figure 4.

Single-cell tumor signatures associated with response to ICI.

a, Volcano plot of expression difference for responder vs. non-responder, PR vs. PD, and PR vs. SD in 12,975 malignant cells from 11 core patients. The log fold change indicates the difference in the mean expression level for each gene. The significance level was determined using two-tailed Wilcoxon Rank Sum test. **b**, Relative sum of loadings for all NMF factors contributed to malignant cells from 11 core patients across RECIST, tissue origins, and cancer subtypes, respectively. **c**, Selection of RECIST-enriched NMF programs. **d**, Enrichment of NMF programs for RECIST groups. Color represents the z-transformed odds ratio. **e**, Expression map of RECIST-specific scINSIGHT modules across individual samples aligned with clinical data. Color represents the z-transformed mean expression of genes contributing to each module. **f**, Enrichment of RECIST-specific gene modules for RECIST groups. Color represents the z-transformed odds ratio. **g**, Hierarchical clustering of pairwise similarities between tumor signatures. INT and UNION, intersection and union of DEGs for responder vs. non-responder, PR vs. PD, and PR vs. SD in Fig. 4a. **h**, Functional categories and **i**, transcription factors of the selected tumor signatures, analyzed by Metascape.

We assessed whether tumor cell signatures are applicable in association with ICI response in other tumors. They had a modest influence on the response to ICI treatment of melanoma (Fig. S8d) in bulk gene expression data [27](#), [28](#).

Combination of tumor signatures and immune cell dynamics classify ICI response

Immune cell dynamics or tumor signature alone has a limited capacity to profile the therapeutic outcome of PD-(L)1 inhibitor alone or in combination (Fig. 5a). Similar to the immune cell blocks and tumor signatures that were over-represented in the poor response group, each of CD4+ Treg, CD4+ TH17, PC7.neg, INT.down, and UNION.down was significantly associated with ICI response in univariate regression analysis (Fig. 5b). The variation in ICI response was not affected by clinical variables of tissue origin, cancer subtype, pathological stage, and smoking status. When we performed a combined analysis of the top tumor-immune features to classify response, the estimation power was improved to over 95% (Fig. 5c). Overall, features of the non-responders, especially CD4+ Treg, B/Plasma cells, INT.down, and UNION.down, showed a higher estimate than those of responders. These non-responder features suggest heterogeneous mechanisms of resistance conferred by tumor and immune regulatory axes.

Discussion

ICI alone, or in combination with chemotherapy, are considered standard first-line therapy for patients with NSCLC. NSCLC may harbor large numbers of genetic perturbations due to genotoxic environmental exposure, which likely generate high mutation burden or neoantigens [29](#). The neoantigen-directed T cell response is hampered by diverse immune suppressive mechanisms exerted by tumor cells and the immune regulatory network [30](#). Current ICIs targeting PD-(L)1 aim one angle of many suppressive mechanisms, and identifying the features of non-responders will reveal additional regulatory angles to improve ICI response.

Immune regulatory network would determine the balance between activation and suppression of tumor-directed immunity. In previous studies, prediction of the response to ICI highlighted CD8+ cytotoxic effector T cells and CD4+ Tregs [31](#). The involvement of effector CD8+ T cells is consistent in most studies regardless of the cellular origin (blood or tissues) or tumor type (melanoma, lung cancer) [32](#). Their phenotypes slightly differ depending on the cellular resolution of the study. Our data provided the highest resolution cell types, and CD8+ TEM cells (*GZMK*, *CXCR4* expression) were overrepresented in the responders. By comparison, Tregs were underrepresented in the responder group. Previously, our group reported a contrasting result demonstrating an increase in Tregs in the posttreatment blood samples (not baseline) in the responder group [33](#). This discrepancy can be explained by the differences in the measurements, sites, and timing of sampling, and the resolution of subpopulations. In mouse preclinical models, PD-L1 inhibitor treatment induces T cell expansion of all phenotypes including CD4+/CD8+ TEFF and Tregs [34](#). Thus, Treg expansion captured in the posttreatment blood samples may represent overall immune activation in human patients. Alternatively, heterogeneity of Tregs and complex effects of PD-(L)1 inhibition on this cell type may contribute to variable results in response prediction.

In search of tumor cell signatures associated with the response to ICI, we adopted two approaches, an individual gene level comparison between the responder and non-responder group and a feature extraction approach to decompose data using the NMF and PCA. Both approaches highlighted attributes of non-responders governed by key transcription factors, which play a significant role in immune response regulation. The ability to predict ICI response based on tumor

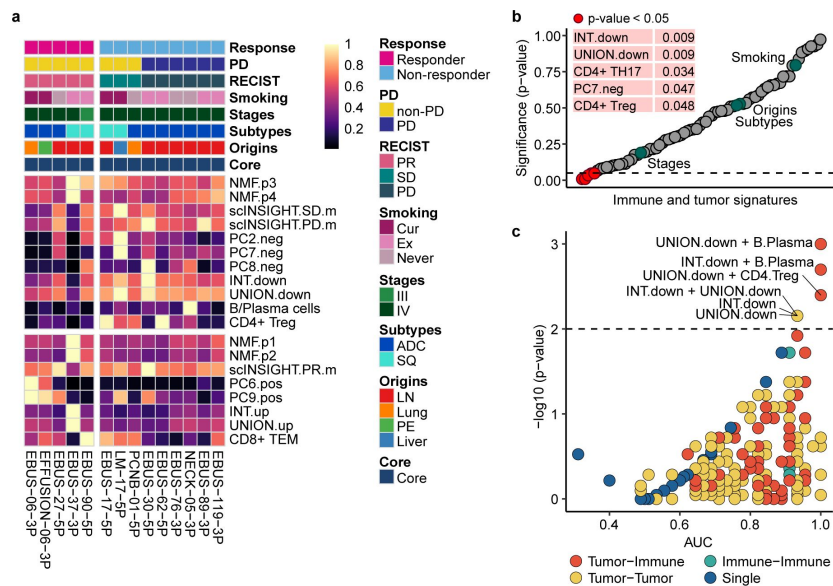


Figure 5.

Combination of tumor signatures and immune index classifying the response to ICI.

a, Heat map of relative contribution of tumor signatures and immune index across individual samples aligned with clinical data. Mean expression of each tumor signature and the percentage of each immune index are divided by the maximum value across samples. INT and UNION, intersection and union of DEGs for responder vs. non-responder, PR vs. PD, and PR vs. SD in **Fig. 4a**. **b**, Univariate regression analysis of immune and tumor signatures for ICI response, together with clinical variables. **c**, Performance of combinatorial index to classify responder and non-responder. p-value, two-tailed Wilcoxon Rank Sum test. P-values adjusted with the Benjamini-Hochberg correction; UNION.down + B/Plasma, 0.22; INT.down + B/Plasma, 0.22; UNION.down + CD4+ Treg, 0.26; INT.down + UNION.down, 0.26; INT.down, 0.26; UNION.down, 0.26.

signatures was as accurate as predictions based on immune cell behavior. Integrating data from both immune and tumor cells enhanced our predictive accuracy, suggesting the presence of both interactive and distinct mechanisms of resistance.

Our study has limitations. Primarily, most samples were obtained from metastatic lymph nodes rather than original tumor tissues, potentially not reflecting the tumor microenvironment accurately. However, prior study [16](#) and Fig. S1 have shown that the immune microenvironment within metastatic lymph nodes closely resembles that of lung tumor tissues, rather than normal lymph nodes. On a positive note, our findings indicate that the immune landscape of metastatic lymph nodes can predict ICI response. Another challenge is the small sample size and the issue of gene expression drop-out, necessitating further studies with a larger patient cohort. Despite these limitations, our study stands out by employing high-throughput scRNA-seq to both tumor cells and immune microenvironment, offering a comprehensive analysis of the multicellular factors that affect ICI response in patients with advanced NSCLC.

Materials and Methods

Human specimens

This study was approved by the Institutional Review Board (IRB) of Samsung Medical Center (IRB no. 2010-04-039-052). Informed written consent was obtained from all patients enrolled in the study. The study participants included 26 patients diagnosed with lung cancer (Table S1). The study population (n=26) has been treated with the investigator's choice either as a clinical trial (n=5) or as standard clinical practice (n=21). Regardless of the treatment selection, the specimens were prospectively collected based on the study protocol. A total of 33 samples were collected and immediately transferred on ice for tissue preparation. Metastatic lymph nodes, metastatic liver tissues, and lung/bronchus tumor tissues from patients with lung cancer were collected using endobronchial ultrasound bronchoscopy, neck lymph node ultrasound and biopsy, liver biopsy, and percutaneous transthoracic cutting needle biopsy. Tumors, normal lungs, normal lymph nodes, and normal brain tissues were obtained during resection surgery. Pleural fluid was collected from patients with malignant pleural effusion.

Clinical outcomes

The clinical outcomes of ICI were evaluated based on the Response Evaluation Criteria in Solid Tumor (RECIST) 1.1 [35](#). In this study, we described non-responders as patients with a stable or progressive disease. Responders were considered as patients with a partial response. None of the patients showed a complete response.

Sample preparation

Single-cell isolation was performed differently depending on the samples. (1) Biopsy samples and normal lymph node tissues were chopped into 2–4 mm pieces and dissociated in an enzyme solution containing collagenase/hyaluronidase (STEMCELL Technologies, Vancouver, Canada) and DNase I, RNase-Free (lyophilized) (QIAGEN, Hilden, Germany) at 37°C for 1 h. Tissue pieces were re-mixed by gentle pipetting at 20-min intervals during incubation. (2) Tumor and normal lung tissue dissociation was performed using a tumor dissociation kit (Miltenyi Biotech, Germany) following the manufacturer's instructions. Briefly, tissue was cut into 2–4 mm pieces and transferred to a C tube containing the enzyme mix (enzymes H, R, and A in RPMI1640 medium). The GentleMACS programs h_tumor_01, h_tumor_02, and h_tumor_02 were run with two 30-min incubations on a MACSmix tube rotator at 37°C. (3) Brain tissue was chopped into 2–4 mm pieces and incubated in an enzyme solution (collagenase (Gibco, Waltham, MA, USA), DNase I (Roche,

Basel, Switzerland), and Dispase I (Gibco) in DMEM) at 37°C for 1 h. Tissue pieces were re-mixed by gentle pipetting at 15-min intervals during incubation. (4) Pleural fluids were transferred to a 50-ml tube, and the cells were spun down at 300g.

Each cell suspension was transferred to a new 50-ml (15-ml for biopsy samples) tube through a 70- μ m strainer. The volume in the tube was readjusted to 50-ml (or 15-ml) with RPMI1640 medium, and spun down to remove the enzymes. The supernatant was aspirated, the cell pellet was resuspended in 4 ml of RPMI1640 medium, and dead cells were removed using Ficoll-Paque PLUS (GE Healthcare, Chicago, IL, USA) separation.

For samples subjected to multiplexing, dissociated cells were cryopreserved in CELLBANKER1 (Zenogen, Fukushima, Japan) and thawed for pooling.

Single-cell RNA sequencing (scRNA-seq) and read processing

Single-cell suspensions were loaded into a Chromium system (10x Genomics, Pleasanton, CA, USA). Following the manufacturer's instructions, 3' scRNA-seq libraries for the 14 samples were generated using Chromium Single Cell 3' v2 Reagent Kits. The 3' library preparation for EBUS_119 used Chromium Single Cell 3' v3 Reagent Kits. The 5' scRNA-seq libraries for twelve individual and two pooled samples were generated using Chromium Single Cell 5' v2 Reagent Kit. Libraries were then sequenced on an Illumina HiSeq 2500 for 3' scRNA-seq and an Illumina NovaSeq 6000 for 5' scRNA-seq. Sequencing reads were mapped to the GRCh38 human reference genome using Cell Ranger toolkit (v5.0.0).

SNP genotyping array

Genomic DNA was extracted from the peripheral blood of six patients and subjected to sample multiplexing (DNeasy Blood & Tissue Kit, QIAGEN). The 766,221 single nucleotide polymorphisms (SNPs) was genotyped using Illumina Global Screening Array MG v2, following the manufacturer's instructions. Normalized signal intensity and genotype were processed using Illumina's GenomeStudio v.2 software.

Demultiplexing of pooled samples

The individuals in sample multiplexing were assigned by a software tool *freemuxlet*, which is an extension of *demuxlet*³⁶ (<https://github.com/statgen/popscl>). First, the *popscl* tool *dsc-pileup* was run with the bam file generated by Cell Ranger toolkit and reference vcf file. The reference was assembled after a lift-over process with GRCh38 from 1000 Genomes Project phase 1 data and the variant allele frequency in East Asian >0.01 were discarded. Next, *freemuxlet* was used to determine the sample identity with default parameters. The individuals were matched based on the similarity between *freemuxlet*-annotated genotypes and SNP array-detected genotypes.

Acquisition of scRNA-seq data from lung adenocarcinoma (LUAD) patients

We obtained raw 3' scRNA-seq from 43 specimens acquired from 33 LUAD patients including early-stage (tLung) and late-stage (tL/B) lung tumor tissues, metastatic lymph nodes (mLN), normal lung tissues (nLung) and lymph nodes (nLN)¹⁶. Sequencing reads were mapped to the GRCh38 human reference genome using Cell Ranger toolkit (v5.0.0).

scRNA-seq data analysis

The raw gene-cell-barcode matrix from Cell Ranger pipeline was processed using Seurat v3.2.2 R package³⁷. Cells were selected using two quality criteria: mitochondrial genes (<20%) and gene count (>200). Cell multiplets predicted by Scrublet³⁸ were filtered out. From the filtered cells, the unique molecular identifier (UMI) count matrix was log-normalized and scaled by z-transform

while regressing out the effects of cell-cycle variations for subsequent analysis. For batch correction, we used Harmony v1.0 R package [39](#) interfacing with Seurat as the *RunHarmony* function. A total of 2,000 variably expressed genes were selected using *FindVariableFeatures* with a parameter `selection.method="vst"`. A subset of principal components (PCs) was selected based on *ElbowPlot* function. Uniform Manifold Approximation and Projection (UMAP) for dimension reduction and cell clustering was performed using *RunUMAP*, *FindNeighbors*, and *FindClusters* functions with the selected PCs and resolutions [Advanced lung cancer patients (Total cells, 33 PCs and resolution=0.3; CD4+ T cells, 28 PCs and resolution=0.9; CD8+ T cells, 28 PCs and resolution=0.9; NK cells, 26 PCs and resolution=0.3; B/Plasma cells, 28 PCs and resolution=0.3; Myeloid cells, 30 PCs and resolution=1.2), LUAD patients (Total cells, 23 PCs and resolution=0.3; T/NK cells, 24 PCs and resolution=0.9)]. We applied the *FindAllMarkers* function to identify differentially expressed genes (DEGs) for each cell cluster. Significance was determined using Wilcoxon Rank Sum test. Genes were selected according to the following statistical thresholds; log fold change>0.25, p-value<0.01, adjusted p-value (bonferroni correction)<0.01, and percentage of cells (pct)>0.25. Cell identity was determined by comparing the expression of known marker genes and DEGs for each cluster.

Principal component analysis (PCA) analysis using the proportion of cell lineages and T/NK cell subsets

PCA analysis was performed for the % proportion of cell lineages and T/NK cell subsets in individual LUAD samples using *prcomp* function of stats v3.6.3 R package. For total cells, the percentages of immune and stromal cells were calculated except for epithelial, cycling, and AMB (ambiguous) cells. For T/NK cells, unknown cells annotated as MT high and AMB cells were excluded.

In silico classification of CD4+ T, CD8+ T, and NK cells

We characterized CD4+ T, CD8+ T, and NK cell populations by combined analysis of gene and protein expression using Cellular Indexing of Transcriptomes and Epitopes by Sequencing data from primary tumor and normal lungs. Among the cells in clusters annotated as T/NK cells, we identified CD3-expressing cells with CD3D or CD3E or CD3G >0 at the RNA level. CD4 and CD8 positive cells were then identified with a cutoff at 55th percentile of antibody-derived tags (ADT) level. NK cells were identified based on the RNA expression level of NK cell markers (*XCL1*, *NCAM1*, *KLRD1*, and *KLRF1*) in CD3 negative cells. The gene expression matrix with cell identity of CD4 positive, CD8 positive, and NK was applied as reference data for supervised cell-type classification using *getFeatureSpace* and *trainModel* functions of scPred v1.9.0 R package [40](#). Finally, we classified T/NK cells in [Fig. 1b](#) into CD4+ T, CD8+ T, and NK cells using scPred *scPredict* function.

Analysis of TCR/BCR repertoires in CD4+ T, CD8+ T, and B/Plasma cells

The data derived from Cell Ranger pipeline for T cell receptor (TCR) and B cell receptor (BCR) sequencing data were processed using scRepertoire v1.2.0 R package [41](#) in R v4.1.1. We selected contigs that generated alpha-beta chain pairs for TCR and heavy-light chain pairs for BCR for subsequent analysis. We called clonotypes based on V(D)J genes and CDR3 nucleotide sequence with the parameter `clonecall="gene+nt"`. The set of clone types was classified by total frequency using the parameter `cloneTypes` defined as Single=1, Small=5, Medium=10, Large=20, and Hyperexpanded=Inf.

Scoring of T cell functional features

Scores for T cell functional features were calculated as the mean expression of regulatory (*ICOS*, *FOXP3*, *IKZF2*, *LAYN*, *TNFRSF18*, *CTLA4*, *IL21R*, *BATF*, *CCR8*, *IL2RA*, and *TNFRSF4*) and cytotoxic (*CX3CR1*, *PRF1*, *GZMA*, *GZMB*, *GZMH*, *GZML*, *GZLN*, *KLRG1*, and *NKG7*) genes at the log-normalized level.

Identification of malignant cells based on inferred copy number variation (CNV) from scRNA-seq data

Two computational tools, inferCNV v1.2.1 (<https://github.com/broadinstitute/inferCNV>) and CopyKAT v1.0.5 R packages⁴², were used to infer genomic copy numbers from scRNA-seq. In a run with inferCNV, the UMI count matrix of each tumor sample was loaded into inferCNV `CreateInfercnvObject` function along with cell lineage annotations. The reference (normal) cells were selected as cells annotated with T/NK, B/Plasma, myeloid, and mast cells. We maintained the proportion of epithelial cells below 20% in each tumor sample using the expression profiles of the normal lung and lymph node tissues. Inferred CNV signals were analyzed using inferCNV `run` function using the parameters: `cutoff=0.1`, `denoise=TRUE`, `HMM=TRUE`, and `HMM_type="i6"`. The signals were then summarized as standard deviations (s.d.) for all windows and the correlation between the CNV in each cell and the mean of the top 5% cells⁴³. Cancer cells showing CNV perturbation (>0.03 s.d. or >0.3 CNV correlation) were classified as malignant cells, otherwise as non-malignant cells. The UMI count matrix of each tumor sample was loaded into CopyKAT `copykat` function along with cell lineage annotations using the following parameters: `ngene.chr=3`, `KS.cut=0.05`, and `norm.cell.names`. Cancer cells predicted as aneuploid cells by CopyKAT were classified as malignant cells. Finally, we identified malignant cells, which are cancer cells classified as malignant cells in either inferCNV or CopyKAT.

Single-cell DEGs between response groups in malignant cells

A total of 12,975 malignant cells were used to identify DEGs in pairwise comparisons according to responder versus non-responder, PR versus PD, and PR versus SD. Differential expression levels were calculated using Seurat `FindMarkers` function with the Wilcoxon Rank Sum test. Genes were selected according to the following statistical thresholds: \log fold change >0.25 , p -value <0.01 , adjusted p -value (bonferroni correction) <0.01 , and $pct >0.25$. We constructed DEGs with intersection (INT) and union (UNION) of up- or down-regulated genes for comparisons.

Non-negative matrix factorization (NMF) programs of the malignant cells

The UMI count matrix for malignant cells was loaded into `nmf` function of RcppML v0.5.6 R package. A NMF model was learned with a rank of 30 using all genes. For each of the 30 NMF factors, the top-ranked 50 genes in the NMF score were defined as signatures. RECIST-enriched NMF program consisted of selected factors based on their relative sum of loadings. We aggregated and redefined gene signatures of factors included in each NMF program. The uniqueness of each NMF program for RECIST groups was evaluated as an odds ratio using `fisher.test` function of stats v3.6.3 R package. Annotations of NMF programs were assigned using Metascape⁴⁴.

RECIST-specific gene modules in malignant cells

The RECIST-specific gene modules were analyzed with a matrix factorization named scINSIGHT²³ using \log -normalized count and 2,000 highly variable genes for each sample. For each module, we selected the 100 genes with the highest coefficients. Combinations of the top 100 genes for modules specific to each RECIST group were defined as module genes. The uniqueness of each module for RECIST groups was evaluated as an odds ratio using `fisher.test` function of stats v3.6.3 R package. Annotations of gene modules were assigned using Metascape⁴⁴.

Principal component signatures of the malignant cells

The UMI count matrix for malignant cells was \log -normalized and scaled by z-transform while regressing out the effects of cell-cycle variations for PCA. A total of 2,000 variably expressed genes selected using `FindVariableFeatures` with `selection.method="vst"` were used for PCA. PCs were

calculated by Seurat *RunPCA* function. PC signatures were selected for 30 genes with + (pos) and – (neg) scores that highly contributed to each PC from PC1 to PC10.

Functional category analysis

Functional categories representing the enriched gene expression in comparisons for responder vs. non-responder, PR vs. PD, and PR vs. SD as well as in the PCs were identified using *fgsea* v1.12.0⁴⁵ R package with parameters: *minSize*=10, *maxSize*=600, and *nperm*=10000. Gene sets for Gene Ontology (GO) Biological Process were collected from the MSigDB database using *msigdb* v7.1.1 R package^{46,47}. The gene list was ranked by the log fold change for each comparison and feature loadings for each PC. Significant GO terms were selected after collapsing redundant terms using *fgsea collapsePathways* function with a statistical threshold for Benjamini-Hochberg adjusted *p*-value<0.05.

Responder-identifying performance tests of PC signatures and combinatorial indexes

Classification models of responders and non-responders for PC signatures and combinatorial indexes between tumor and/or immune cells were tested by receiver operating characteristic (ROC) curve. Relative numbers between the observed and expected cells (*Ro/e*) for each sample were obtained from the chi-square test¹⁸. Area under the curve (AUC) was calculated using *ROC* function of *Epi* v2.44 R package with *Ro/e* scores as input. Significance was calculated by Wilcoxon Rank Sum test and confirmed by adjusting with the Benjamini-Hochberg correction.

Univariate regression analysis for ICI response

Univariate regression was performed using the *lm* function of *stats* v3.6.3 with *Ro/e* scores for each sample as input. We evaluated the relationship between the target variable ICI response, classified as responders and non-responders, and one predictor variable of immune cell types, tumor signatures, and clinical factors such as tissue origin, cancer subtype, pathological stage, and smoking status. The significance of predictor was calculated using *Anova* function of *car* v3.0-9 R package.

Validation of PC signatures in melanoma cohorts

We used Riaz et al.'s²⁸ and Van Allen et al.'s²⁷ RNA sequencing data from melanoma patients receiving PD-1 and CTLA-4 immune checkpoint therapy to assess expressional changes of PC signatures along RECIST. The mean expression of each PC signature in each RECIST group was calculated as the log₂ normalized level.

Data availability

Raw single-cell RNA sequencing data generated during the current study are available in the European Genome-phenome Archive (EGA) database (accession code EGAD00001008703), and processed data can be accessed from the NCBI Gene Expression Omnibus (GEO) database (accession code GSE205335; temporary token 'khatsauybfcljgb'). Single-cell RNA sequencing data for LUAD patients analyzed in this study are available in the EGA database at EGAD00001005054. Bulk RNA sequencing data analyzed in this study were obtained from GEO at GSE91061 and database of Genotypes and Phenotypes (dbGap) at phs000452.v2.p1.

Acknowledgements

This study was supported by the Collaborative Genome Program for Fostering New Post-Genome Industry (NRF-2017M3C9A6044633 and NRF-2017M3C9A6044636), Mid-Career Researcher Program (NRF-2022R1A2C1091451), and Basic Research Laboratory Program (RS-2023-00220840) of the National Research Foundation of Korea funded by the Korea government. We also acknowledge the Basic Medical Science Facilitation Program, through the Catholic Medical Center of the Catholic University of Korea funded by the Catholic Education Foundation and the KREONET/GLORIAD service provided by KISTI (Korea Institute of Science and Technology Information).

Supplementary Files

Supplementary file 1: Fig. S1-S8. Supplementary figures and legends.

Supplementary file 2: Table S1. Patient information of lung cancer ICI cohorts.

Supplementary file 3: Table S2. List of genes specific to the cell clusters in total cells and each cell lineage.

Supplementary file 4: Table S3. List of DEGs in comparison between ICI response groups.

Supplementary file 5: Table S4. Details of GO terms significantly enriched in comparisons for ICI response groups and PCs.

Supplementary file 6: Table S5. Gene list of tumor signatures.

References

- 1 Reck M., et al. (2016) **Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer** *N Engl J Med* **375**:1823–1833 <https://doi.org/10.1056/NEJMoa1606774>
- 2 Gandhi L., et al. (2018) **Pembrolizumab plus Chemotherapy in Metastatic Non-Small-Cell Lung Cancer** *N Engl J Med* **378**:2078–2092 <https://doi.org/10.1056/NEJMoa1801005>
- 3 Paz-Ares L., et al. (2018) **Pembrolizumab plus Chemotherapy for Squamous Non-Small-Cell Lung Cancer** *N Engl J Med* **379**:2040–2051 <https://doi.org/10.1056/NEJMoa1810865>
- 4 Wakelee H A. N. *et al.* (2021) **IMpower010: Primary results of a phase III global study of atezolizumab versus best supportive care after adjuvant chemotherapy in resected stage IB-III A non-small cell lung cancer (NSCLC)** *J Clin Oncol* **39** https://doi.org/10.1200/JCO.2021.39.15_suppl.8500
- 5 Reck M., et al. (2021) **Five-Year Outcomes With Pembrolizumab Versus Chemotherapy for Metastatic Non-Small-Cell Lung Cancer With PD-L1 Tumor Proportion Score \geq 50** *J Clin Oncol* **39**:2339–2349 <https://doi.org/10.1200/JCO.21.00174>

- 6 Fehrenbacher L., et al. (2016) **Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial** *Lancet* **387**:1837–1846 [https://doi.org/10.1016/S0140-6736\(16\)00587-0](https://doi.org/10.1016/S0140-6736(16)00587-0)
- 7 Hellmann M. D., et al. (2018) **Nivolumab plus Ipilimumab in Lung Cancer with a High Tumor Mutational Burden** *N Engl J Med* **378**:2093–2104 <https://doi.org/10.1056/NEJMoa1801946>
- 8 Litchfield K., et al. (2021) **Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition** *Cell* **184**:596–614 <https://doi.org/10.1016/j.cell.2021.01.002>
- 9 Kamphorst A. O., et al. (2017) **Proliferation of PD-1+ CD8 T cells in peripheral blood after PD-1-targeted therapy in lung cancer patients** *Proc Natl Acad Sci U S A* **114**:4993–4998 <https://doi.org/10.1073/pnas.1705327114>
- 10 Kim K. H., et al. (2019) **The First-week Proliferative Response of Peripheral Blood PD-1(+)/CD8(+) T Cells Predicts the Response to Anti-PD-1 Therapy in Solid Tumors** *Clin Cancer Res* **25**:2144–2154 <https://doi.org/10.1158/1078-0432.CCR-18-1449>
- 11 Thommen D. S., et al. (2018) **A transcriptionally and functionally distinct PD-1(+)/CD8(+) T cell pool with predictive potential in non-small-cell lung cancer treated with PD-1 blockade** *Nat Med* **24**:994–1004 <https://doi.org/10.1038/s41591-018-0057-z>
- 12 Arce Vargas F., et al. (2017) **Fc-Optimized Anti-CD25 Depletes Tumor-Infiltrating Regulatory T Cells and Synergizes with PD-1 Blockade to Eradicate Established Tumors** *Immunity* **46**:577–586 <https://doi.org/10.1016/j.immuni.2017.03.013>
- 13 Krieg C., et al. (2018) **Author Correction: High-dimensional single-cell analysis predicts response to anti-PD-1 immunotherapy** *Nat Med* **24**:1773–1775 <https://doi.org/10.1038/s41591-018-0094-7>
- 14 Kumagai S., et al. (2020) **The PD-1 expression balance between effector and regulatory T cells predicts the clinical efficacy of PD-1 blockade therapies** *Nat Immunol* **21**:1346–1358 <https://doi.org/10.1038/s41590-020-0769-3>
- 15 Ruiz-Patino A., et al. (2020) **Immunotherapy at any line of treatment improves survival in patients with advanced metastatic non-small cell lung cancer (NSCLC) compared with chemotherapy (Quijote-CLICaP)** *Thorac Cancer* **11**:353–361 <https://doi.org/10.1111/1759-7714.13272>
- 16 Kim N., et al. (2020) **Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma** *Nat Commun* **11** <https://doi.org/10.1038/s41467-020-16164-1>
- 17 Stoeckius M., et al. (2017) **Simultaneous epitope and transcriptome measurement in single cells** *Nat Methods* **14**:865–868 <https://doi.org/10.1038/nmeth.4380>
- 18 Guo X., et al. (2018) **Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing** *Nat Med* **24**:978–985 <https://doi.org/10.1038/s41591-018-0045-3>
- 19 Gueguen P., et al. (2021) **Contribution of resident and circulating precursors to tumor-infiltrating CD8(+) T cell populations in lung cancer** *Sci Immunol* **6** <https://doi.org/10.1126/sciimmunol.abd5778>

- 20 Groom J. R., Luster A. D (2011) **CXCR3 in T cell function** *Exp Cell Res* **317**:620–631 <https://doi.org/10.1016/j.yexcr.2010.12.017>
- 21 Yang C., et al. (2019) **Heterogeneity of human bone marrow and blood natural killer cells defined by single-cell transcriptome** *Nat Commun* **10** <https://doi.org/10.1038/s41467-019-11947-7>
- 22 Liu L., et al. (2019) **Combination of TMB and CNA Stratifies Prognostic and Predictive Responses to Immunotherapy Across Metastatic Cancer** *Clin Cancer Res* **25**:7413–7423 <https://doi.org/10.1158/1078-0432.CCR-19-0558>
- 23 Qian K., Fu S., Li H., Li W. V (2022) **scINSIGHT for interpreting single-cell gene expression from biologically heterogeneous data** *Genome Biol* **23** <https://doi.org/10.1186/s13059-022-02649-3>
- 24 Lapuente-Santana O., van Genderen M., Hilbers P. A. J., Finotello F., Eduati F (2021) **Interpretable systems biomarkers predict response to immune-checkpoint inhibitors** *Patterns (N Y)* **2** <https://doi.org/10.1016/j.patter.2021.100293>
- 25 Betzler A. C., et al. (2020) **NF-kappaB and Its Role in Checkpoint Control** *Int J Mol Sci* **21** <https://doi.org/10.3390/ijms21113949>
- 26 Fenton S. E., Saleiro D., Platanius L. C (2021) **Type I and II Interferons in the Anti-Tumor Immune Response** *Cancers (Basel)* **13** <https://doi.org/10.3390/cancers13051037>
- 27 Van Allen E. M., et al. (2015) **Genomic correlates of response to CTLA-4 blockade in metastatic melanoma** *Science* **350**:207–211 <https://doi.org/10.1126/science.aad0095>
- 28 Riaz N., et al. (2017) **Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab** *Cell* **171**:934–949 <https://doi.org/10.1016/j.cell.2017.09.028>
- 29 Lawrence M. S., et al. (2013) **Mutational heterogeneity in cancer and the search for new cancer-associated genes** *Nature* **499**:214–218 <https://doi.org/10.1038/nature12213>
- 30 Sharma P., Hu-Lieskovan S., Wargo J. A., Ribas A. (2017) **Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy** *Cell* **168**:707–723 <https://doi.org/10.1016/j.cell.2017.01.017>
- 31 Gibellini L., et al. (2020) **Single-Cell Approaches to Profile the Response to Immune Checkpoint Inhibitors** *Front Immunol* **11** <https://doi.org/10.3389/fimmu.2020.00490>
- 32 Zheng L., et al. (2021) **Pan-cancer single-cell landscape of tumor-infiltrating T cells** *Science* **374** <https://doi.org/10.1126/science.abe6474>
- 33 Koh J., et al. (2020) **Regulatory (FoxP3(+)) T cells and TGF-beta predict the response to anti-PD-1 immunotherapy in patients with non-small cell lung cancer** *Sci Rep* **10** <https://doi.org/10.1038/s41598-020-76130-1>
- 34 Wei S. C., et al. (2019) **Combination anti-CTLA-4 plus anti-PD-1 checkpoint blockade utilizes cellular mechanisms partially distinct from monotherapies** *Proc Natl Acad Sci U S A* **116**:22699–22709 <https://doi.org/10.1073/pnas.1821218116>

- 35 Eisenhauer E. A., et al. (2009) **New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1)** *Eur J Cancer* **45**:228–247 <https://doi.org/10.1016/j.ejca.2008.10.026>
- 36 Kang H. M., et al. (2018) **Multiplexed droplet single-cell RNA-sequencing using natural genetic variation** *Nat Biotechnol* **36**:89–94 <https://doi.org/10.1038/nbt.4042>
- 37 Stuart T., et al. (2019) **Comprehensive Integration of Single-Cell Data** *Cell* **177**:1888–1902 <https://doi.org/10.1016/j.cell.2019.05.031>
- 38 Wolock S. L., Lopez R., Klein A. M (2019) **Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data** *Cell Syst* **8**:281–291 <https://doi.org/10.1016/j.cels.2018.11.005>
- 39 Korsunsky I., et al. (2019) **Fast, sensitive and accurate integration of single-cell data with Harmony** *Nat Methods* **16**:1289–1296 <https://doi.org/10.1038/s41592-019-0619-0>
- 40 Alquicira-Hernandez J., Sathe A., Ji H. P., Nguyen Q., Powell J. E (2019) **scPred: accurate supervised method for cell-type classification from single-cell RNA-seq data** *Genome Biol* **20** <https://doi.org/10.1186/s13059-019-1862-5>
- 41 Borcherding N., Bormann N. L., Kraus G (2020) **scRepertoire: An R-based toolkit for single-cell immune receptor analysis** *F1000Res* **9** <https://doi.org/10.12688/f1000research.22139.2>
- 42 Gao R., et al. (2021) **Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes** *Nat Biotechnol* **39**:599–608 <https://doi.org/10.1038/s41587-020-00795-2>
- 43 Puram S. V., et al. (2017) **Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer** *Cell* **171**:1611–1624 <https://doi.org/10.1016/j.cell.2017.10.044>
- 44 Zhou Y., et al. (2019) **Metascape provides a biologist-oriented resource for the analysis of systems-level datasets** *Nat Commun* **10** <https://doi.org/10.1038/s41467-019-09234-6>
- 45 Korotkevich G, Sergushichev A. (2019) **Fast gene set enrichment analysis** *bioRxiv* <https://doi.org/10.1101/060012>
- 46 Liberzon A., et al. (2011) **Molecular signatures database (MSigDB) 3.0** *Bioinformatics* **27**:1739–1740 <https://doi.org/10.1093/bioinformatics/btr260>
- 47 Subramanian A., et al. (2005) **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles** *Proc Natl Acad Sci U S A* **102**:15545–15550 <https://doi.org/10.1073/pnas.0506580102>

Editors

Reviewing Editor

Ching-Hao Wang

GlaxoSmithKline, Cambridge, United States of America

Senior Editor

Tony Ng

King's College London, London, United Kingdom

Reviewer #1 (Public Review):

Summary:

The authors study the variability of patient response of NSCLC patients on immune checkpoint inhibitors using single-cell RNA sequencing in a cohort of 26 patients and 33 samples (primary and metastatic sites), mainly focusing on 11 patients and 14 samples for association analyses, to understand the variability of patient response based on immune cell fractions and tumor cell expression patterns. The authors find immune cell fraction, clonal expansion differences, and tumor expression differences between responders and non-responders. Integrating immune and tumor sources of signal the authors claim to improve prediction of response markedly, albeit in a small cohort.

Strengths:

- The problem of studying the tumor microenvironment, as well as the interplay between tumor and immune features is important and interesting and needed to explain the heterogeneity of patient response and be able to predict it.
- Extensive analysis of the scRNAseq data with respect to immune and tumor features on different axes of hypothesis relating to immune response and tumor immune evasion using state-of-the-art methods.
- The authors provide an interesting scRNAseq data set linked to outcomes data.
- Integration of TCRseq to confirm subtype of T-cell annotation and clonality analysis.
- Interesting analysis of cell programs/states of the (predicted) tumor cells and characterization thereof.

Weaknesses:

- Generally, a very heterogeneous and small cohort where adjustments for confounding are hard. Additionally, there are many tests for association with outcome, where necessary multiple testing adjustments would negate signal and confirmation bias likely, so biological takeaways have to be questioned.
- RNAseq is heavily influenced by the tissue of origin (both cell type and expression), so the association with the outcome can be confounded. The authors try to argue that lymph node T-cell and NK content are similar, but a quantitative test on that would be helpful.
- The authors claim a very high "accuracy" performance, however, given the small cohort and lack of information on the exact evaluation it is not clear if this just amounts to overfitting the data.
- Especially for tumor cell program/state analysis the specificity to the setting of ICIs is not clear and could be prognostic.
- Due to the small cohort with a lot of variability, more external validation is needed to be convincingly reproducible, especially when talking about AUC/accuracy of a predictor.

<https://doi.org/10.7554/eLife.98366.1.sa1>

Reviewer #2 (Public Review):

Summary:

The authors have utilised deep profiling methods to generate deeper insights into the features of the TME that drive responsiveness to PD-1 therapy in NSCLC.

Strengths:

The main strengths of this work lie in the methodology of integrating single-cell sequencing, genetic data, and TCRseq data to generate hypotheses regarding determinants of IO responsiveness.

Some of the findings in this study are not surprising and well preceded eg. association of Treg, STAT3, and NFkB with ICI resistance and CD8⁺ activation in ICI responders and thus act as an additional dataset to add weight to this prior body of evidence. Whilst the role of Th17 in PD-1 resistance has been previously reported (eg. *Cancer Immunol Immunother* 2023 Apr;72(4):1047-1058, *Cancer Immunol Immunother* 2024 Feb 13;73(3):47, *Nat Commun.* 2021; 12: 2606) these studies have used non-clinical models or peripheral blood readouts. Here the authors have supplemented current knowledge by characterization of the TME of the tumor itself.

Weaknesses:

Unfortunately, the study is hampered by the small sample size and heterogeneous population and whilst the authors have attempted to bring in an additional dataset to demonstrate the robustness of their approach, the small sample size has limited their ability to draw statistically supported conclusions. There is also limited validation of signatures/methods in independent cohorts, no functional characterisation of the findings, and the discussion section does not include discussion around the relevance/interpretation of key findings that were highlighted in the abstract (eg. role of Th17, TRM, STAT3, and NFkB). Because of these factors, this work (as it stands) does have value to the field but will likely have a relatively low overall impact.

Related to the absence of discussion around prior TRM findings, the association between TRM involvement in response to IO therapy in this manuscript is counter to what has been previously demonstrated (*Cell Rep Med.* 2020;1(7):100127, *Nat Immunol.* 2017;18(8):940-950., *J Immunol.* 2015;194(7):3475-3486.). However, it should be noted that the authors in this manuscript chose to employ alternative markers of TRM characterisation when defining their clusters and this could indicate a potential rationale for differences in these findings. TRM population is generally characterised through the inclusion of the classical TRM markers CD69 (tissue retention marker) and CD103 (TCR experienced integrin that supports epithelial adhesion), which are both absent from the TRM definition in this study. Additional markers often used are CD44, CXCR6, and CD49a, of which only CXCR6 has been included by the authors. Conversely, the majority of markers used by the authors in the cell type clustering are not specific to TRM (eg. CD6, which is included in the TRM cluster but is expressed at its lowest in cluster 3 which the authors have highlighted as the CD8⁺ TRM population). Therefore, whilst there is an interesting finding of this particular cell cluster being associated with resistance to ICI, its annotation as a TRM cluster should be interpreted with caution.

<https://doi.org/10.7554/eLife.98366.1.sa0>

Author response:

We appreciate the comprehensive reviews and would like to address the critiques and suggestions provided by both reviewers. We will make significant revisions to the manuscript to address these concerns. These include a more cautious interpretation of our results, an expanded discussion on key findings, additional analyses for TRM characterization, and a

clearer outline of future validation efforts. We believe these changes will enhance the clarity and robustness of our study, and we hope they meet the reviewer's expectations.

Reviewer 1:

Weaknesses:

(1) Heterogeneous and small cohort:

Increasing the cohort size is not feasible due to resource constraints. We acknowledge the challenges posed by the heterogeneous and small cohort, which complicate adjustments for confounding. We will apply multiple testing corrections to transparently assess and accurately report the robustness of our findings in the revision.

(2) Influence of tissue of origin on RNAseq:

We agree that RNAseq results can be heavily influenced by the tissue of origin. While immune cell composition in the normal lung tissues and lymph nodes is quite different, we found that in tumor tissues and metastatic lymph nodes, these differences diminish and common features dominate. Although we depicted this data in the supplementary figure 1, we did not provide a quantitative test in the original submission. In the revision, we will perform additional quantitative tests to compare immune cell composition across different tissue origins. These tests will provide a more precise understanding of the cellular composition and support our argument regarding the similarity of tumor-sculpted microenvironment. We will include these results and detailed methodologies in the revision.

(3) Accuracy performance and overfitting:

We acknowledge the concern regarding the high "accuracy" performance potentially indicating overfitting. We will clarify the evaluation methods used and moderate our claims regarding accuracy in the revision.

(4) Specificity of the tumor cell program/state analysis to the setting of ICIs:

The comment suggests that the tumor programs in our study may not be specific to the ICI group but rather prognostic in lung cancer. We acknowledge this possibility as we performed comparisons between responders and non-responders (with different cut-offs) to find common trends and interpreted them in terms of their association with ICI. In the revision, we will test the prognostic association of the tumor programs using public lung cancer data.

(5) More external validation needed:

We recognize the importance of external validation for reproducibility. While increasing the cohort size is not feasible, we will propose future directions for validation using larger, independent cohorts and potential experimental validations.

Reviewer 2:

Weaknesses:

(1) Small sample size and heterogeneous populations:

Increasing the cohort size is not feasible due to resource constraints. We acknowledge the challenges posed by the heterogeneous and small cohort, which complicate adjustments for

confounding. We will apply multiple testing corrections to transparently assess and accurately report the robustness of our findings in the revision.

| (2) *Limited validation of signatures/ methods in independent cohorts:*

We recognize the importance of external validation for reproducibility. While increasing the cohort size is not feasible, we will propose future directions for validation using larger, independent cohorts and potential experimental validations.

| (3) *Lack of functional characterization and discussion on key findings:*

We appreciate the feedback regarding the need for functional characterization and a more thorough discussion of key findings on the roles of specific cell populations and genes. In the revised manuscript, we will expand the discussion section to include in-depth analysis of these findings and their relevance to the study. This includes a detailed interpretation of how these factors contribute to the immune response and potential implications for therapy.

| (4) *TRM findings and marker selection:*

We understand the concern regarding the association between TRM involvement in response to IO therapy, which appears counter to previous demonstrations. It is indeed important to note that we employed alternative markers for TRM characterization. Our choice of markers was based on transcriptional references relevant to our study. However, we agree that classical TRM markers such as CD69 and CD103, which were absent in our definition, are critical for accurate TRM identification. To address this, we will include a detailed rationale for our marker selection and acknowledge the limitations of our TRM characterization. We will include additional analyses using classical TRM markers where possible and incorporate these findings into the revision. This will provide a clearer understanding of our TRM population and its role in the immune response to IO therapy.

<https://doi.org/10.7554/eLife.98366.1.sa3>