Journal for ImmunoTherapy of Cancer

Spatially resolved proteomic profiling identifies tumor cell CD44 as a biomarker associated with sensitivity to PD-1 axis blockade in advanced non-small-cell lung cancer

Myrto K Moutafi ,[1] Magdalena Molero,[2,3] Sandra Martinez Morilla,[1] Javier Baena,[2,4] Ioannis A Vathiotis ,[1] Niki Gavrielatou ,[1] Laura Castro-Labrador,[2] Gorka Ruiz de Garibay ,[2,3] Vera Adradas,[2] Daniel Orive,[5,6] Karmele Valencia,[5,7,8] Alfonso Calvo ,[5,6,7,8] Luis M Montuenga,[5,6,7,8] S Ponce Aix,[2,3,4,7] Kurt A Schalper ,[1,9] Roy S Herbst,[9] Luis Paz-Ares,[2,3,4,7,10] David L Rimm ,[1,9] Jon Zugazagoitia [2,3,4,7]

To cite: Moutafi MK, Molero M, Martinez Morilla S, *et al*. Spatially resolved proteomic profiling identifies tumor cell CD44 as a biomarker associated with sensitivity to PD-1 axis blockade in advanced non-small-cell lung cancer. *Journal for ImmunoTherapy of Cancer* 2022;**10**:e004757. doi:10.1136/jitc-2022-004757

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/jitc-2022-004757).

MKM and MM contributed equally.

Accepted 17 July 2022

For numbered affiliations see end of article.

Correspondence to
Dr Jon Zugazagoitia;
j.zugazagoitia.imas12@h12o.es

ABSTRACT

Background Most patients with advanced non-small-cell lung cancer (NSCLC) fail to derive significant benefit from programmed cell death protein-1 (PD-1) axis blockade, and new biomarkers of response are needed. In this study, we aimed to discover and validate spatially resolved protein markers associated with sensitivity to PD-1 axis inhibition in NSCLC.

Methods We initially assessed a discovery cohort of 56 patients with NSCLC treated with PD-1 axis inhibitors at Yale Cancer Center. Using the GeoMx Digital Spatial Profiling (DSP) system, 71 proteins were measured in spatial context on each spot in a tissue microarray. We used the AQUA method of quantitative immunofluorescence (QIF) to orthogonally validate candidate biomarkers. For external independent validation, we assessed whole tissue sections derived from 128 patients with NSCLC treated with single-agent PD-1 axis inhibitors at the 12 de Octubre Hospital (Madrid) using DSP. We further analyzed two immunotherapy untreated cohorts to address prognostic significance (n=252 from Yale Cancer Center; n=124 from University Clinic of Navarra) using QIF and DSP, respectively.

Results Using continuous log-scaled data, we identified CD44 expression in the tumor compartment (pan-cytokeratin (CK)+) as a novel predictor of prolonged progression-free survival (PFS) (multivariate HR=0.68, p=0.043) in the discovery set. We validated by QIF that tumor CD44 levels assessed as continuous QIF scores were associated with longer PFS (multivariate HR=0.31, p=0.022) and overall survival (multivariate HR=0.29, p=0.038). Using DSP in an independent immunotherapy treated cohort, we validated that CD44 levels in the tumor compartment, but not in the immune compartment (panCK–/CD45+), were associated with clinical benefit (OR=1.22, p=0.018) and extended PFS under PD-1 axis inhibition using the highest tertile cutpoint (multivariate HR=0.62, p=0.03). The effect of tumor cell CD44 in predicting PFS remained significant after correcting for programmed death-ligand 1 (PD-L1) Tumor Proportion Score (TPS) in both cohorts. High tumor cell CD44 was not prognostic in the absence of immunotherapy. Using DSP data, intratumoral regions with elevated tumor cell CD44 expression showed prominent (fold change>1.5, adjusted p<0.05) upregulation of PD-L1, TIM-3, ICOS, and CD40 in two independent cohorts.

Conclusions This work highlights CD44 as a novel indicative biomarker of sensitivity to PD-1 axis blockade that might help to improve immunotherapy strategies for NSCLC.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ CD44 has been identified as a programmed death-ligand 1 (PD-L1) regulator at both the protein and the messenger RNA levels, and it has been shown to contribute to PD-L1-mediated T cell suppression in preclinical models. However, the association between tumor cell CD44 overexpression and improved clinical outcomes to programmed cell death protein-1 (PD-1) axis blockade has not been previously reported.

WHAT THIS STUDY ADDS

⇒ This study identifies and validates that CD44, only when expressed by tumor cells and not immune cells, is a marker of sensitivity to PD-1 blockade that is non-redundant with PD-L1 expression.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ CD44 could complement existing biomarkers for optimal patient stratification, and potentially open new therapeutic strategies to improve precision immunotherapy in lung cancer. Further mechanistic studies are needed to understand the interplay between CD44 positive cancer stem cell phenotype and mechanisms of immune evasion.

INTRODUCTION

Programmed cell death protein-1 (PD-1) axis blockade has transformed the

treatment landscape of advanced non-small cell lung cancer (NSCLC). Antibodies blocking the PD-1/programmed death-ligand 1 (PD-L1) interaction have shown improved survival as single agents, in combination with chemotherapy, or combined with cytotoxic T-lymphocytes-associated protein 4 (CTLA-4) blockade.[1 2] However, most patients fail to develop durable responses particularly when these drugs are given as monotherapies in unselected patients.[1 2] Moreover, there is a limited understanding of the mechanisms of sensitivity and resistance to PD-1 axis inhibitors, and robust biomarkers of response are also lacking.[2 3]

The characterization of tumor cells and their surrounding microenvironment by direct analysis of tumor specimens has shown utility to predict outcomes from PD-1 axis blockade. Notably, in a meta-analysis involving 10 different tumor types and more than 8000 tumor specimens, multiplexed immunohistochemistry (IHC)/immunofluorescence showed significantly higher performance than PD-L1 expression (assessed by IHC), tumor mutational burden, or bulk interferon (IFN)-γ-based gene expression signatures for discriminating between responders and non-responders to immune checkpoint blockade.[4] These findings underscore the importance of assessing immunotherapy biomarkers in a quantitative manner and in their spatial context, and illustrate the potential of these technologies to identify new mechanisms of response to immunotherapy.

CD44 is a transmembrane glycoprotein expressed as a wide variety of isoforms in most cells of the human body,[5] and has been implicated in multiple pathways that are essential for tumor maintenance and progression. Most notably, CD44 participates in key aspects of stemness,[6 7] particularly epithelial to mesenchymal transition (EMT),[8–11] migration,[9 11 12] and apoptosis resistance.[5] Beyond its roles in regulating tumor-intrinsic activities, CD44 can also modulate antitumor immunity. For instance, CD44 has been identified as a PD-L1 regulator at both the protein and the messenger RNA (mRNA) levels, and it has been shown to contribute to PD-L1-mediated T cell suppression in preclinical models.[13 14] However, the distribution and patterns of CD44 protein expression and its potential role in promoting sensitivity to PD-1 axis blockade in human NSCLC has not been reported.

In this study, we aimed to discover and validate spatially resolved protein markers associated with sensitivity to PD-1 axis inhibition in advanced NSCLC. Using two orthogonal spatial proteomic technologies in immunotherapy treated and immunotherapy untreated NSCLC cohorts from different institutions, we identify and validate tumor cell CD44 as a novel indicative biomarker of sensitivity to PD-1 axis blockade. In addition, we show that intratumoral regions with elevated expression of CD44 in the tumor compartment display a unique immune microenvironment, characterized by prominent upregulation of multiple immunomodulatory molecules.

METHODS

Patient cohorts

We analyzed retrospectively collected, formalin-fixed, paraffin-embedded tumor specimens from four independent multi-institutional NSCLC cohorts.

We used two immunotherapy treated NSCLC cohorts as discovery and validation sets. The discovery set (YTMA471) contained 56 patients that received immune checkpoint blockade for advanced disease between 2012 and 2019 at Yale Cancer Center (New Haven, Connecticut, USA), whose pre-treatment tumors were represented in a tissue microarray (TMA) format. The validation cohort (H12O_ITX1) was composed with 128 patients that were treated with single-agent PD-1 axis inhibitors between 2013 and 2019 at the 12 de Octubre University Hospital (Madrid, Spain), whose pre-treatment tumors were assessed as whole tissue sections. The baseline characteristics of these cohorts are summarized in table 1.

To rule out prognostic significance of the validated predictors, we analyzed two control cohorts of immunotherapy untreated patients with NSCLC, whose tumors were represented in TMA formats: YTMA423 (Yale control cohort), which contained 252 evaluable tumors resected between 2011 and 2016 at Yale Cancer Center; and University Clinic of Navarra (CUN) cohort (CIMA-CUN control cohort), containing 124 evaluable tumors resected between 2000 and 2013 at CUN (Pamplona, Spain). The baseline characteristics of the two control cohorts are summarized in online supplemental table S1. The details for TMA construction are summarized in the online supplemental materials and methods.

Digital spatial profiling

Tissue slides from YTMA471, H12O_ITX1, and CIMA-CUN cohorts were subjected to digital spatial profiling (DSP) experiments. Two slides derived from independent YTMA471 blocks, each block containing two non-adjacent tumor cores per NSCLC case, were profiled at Yale (Rimm Lab). One hundred and twenty-eight whole tissue sections from H12O_ITX1, and 6 TMA slides from CIMA-CUN cohort, each containing three tumor cores per NSCLC case, were profiled at the 12 de Octubre Research Facilities (Zugazagoitia Lab).

Briefly, the slides were first deparaffinized and subjected to antigen retrieval procedures, then we co-incubated them with fluorescent labeled antibodies (morphology markers), together with photocleavable oligonucleotide-labeled primary antibodies (profiling antibodies). In YTMA471 cohort, we used three morphology markers to detect tumor cells (pan-cytokeratin (CK)), all immune cells (CD45), and macrophages (CD68), together with a 71-plex profiling antibody panel. In cohorts H12O_ITX1 and CIMA-CUN, we used two morphology markers (panCK and CD45), together with a 36-plex profiling antibody panel (online supplemental table S2). Once the staining step was completed, we loaded the slides on the GeoMx DSP instruments (NanoString), where they were scanned to produce a digital fluorescent image of

Table 1 Baseline patient characteristics of the immunotherapy treated NSCLC cohorts

Characteristic	YTMA471 cohort N (%)	H12O_ITX1 cohort N (%)
Total	56	128
Gender		
Male	28 (50)	100 (78.1)
Female	28 (50)	28 (21.9)
Age		
<70 years old	30 (53.6)	68 (53.1)
≥70 years old	26 (46.4)	60 (46.9)
Smoking history		
Active smoker	10 (18)	34 (26.6)
Former smoker	44 (79)	88 (68.7)
Never smoker	2 (3.6)	6 (4.7)
Histology		
Adenocarcinoma	41 (73)	54 (42.2)
Large-cell carcinoma	–	16 (12.5)
Squamous cell carcinoma	9 (16)	58 (45.3)
Adenosquamous carcinoma	2 (3.6)	–
NSCLC-NOS	4 (7.1)	–
Stage		
III	5 (8.9)	5 (3.9)
IV	51 (91.1)	123 (96.1)
Actionable drivers		
EGFR/ALK/ROS1/RET alterations	1 (1.7)	7 (5.4)
No *EGFR/ALK/ROS1/RET* alterations	48 (85.7)	63 (49.2)
Unknown genetic alterations	7 (12.5)	58 (45.3)
Central nervous system metastasis		
Yes	21 (38)	24 (18.7)
No	35 (62)	104 (81.3)
Liver metastasis		
Yes	11 (20)	25 (19.5)
No	45 (80)	103 (80.5)
LIPI score		
Good	9 (16.1)	27 (21.1)
Intermediate	22 (39.3)	54 (42.2)
Poor	3 (5.4)	42 (32.8)
Unknown	22 (39.3)	5 (3.9)
PD-L1 TPS by IHC		
≥50%	15 (26.8)	32 (25)
1%–49%	20 (35.7)	44 (34.4)
<1%	11 (19.6)	46 (36)
Not evaluable	10 (17.9)	6 (4.6)
Immunotherapy line		
1	39 (70)	20 (15.6)
2	14 (25)	79 (61.7)
≥3	3 (5.4)	29 (12.5)
Type of immunotherapy		
Chemotherapy +PD-1 axis blockade	22 (39.3)	0
Dual PD-1 +CTLA4 blockade	1 (1.8)	0
Other PD-1 axis-based combinations	5 (8.9)	0
Single-agent PD-1 axis blockade	28 (50)	128 (100)
Best response to immunotherapy		
Complete or partial response	17 (33.5)	22 (17.1)
Stable disease	22 (39)	32 (25)
Progressive disease	15 (27)	74 (57.8)
Not evaluable	1	0
Clinical benefit group		
Clinical benefit	36 (64.3)	37 (28.9)
No clinical benefit	19 (33.9)	90 (70.3)
Not evaluable	1 (1.8)	1 (0.8)

CTLA-4, cytotoxic T-lymphocytes-associated protein 4 ; IHC, immunohistochemistry; LIPI, Lung Immune Prognostic Index; NSCLC-NOS, non-small cell lung cancer not otherwise specified; PD-1, programmed cell death protein-1 ; PD-L1, programmed death-ligand 1 ; TPS, Tumor Proportion Score.

the tissue. Next, we generated circular regions of interest (ROIs) of a maximum of 660 µm in each slide. In TMA slides, ROIs were restricted to the tissue cores (generally one ROI per core was generated) (figure 1). For whole tissue slides, we selected multiple ROIs from intratumoral immune-enriched areas (panCK+/CD45+) distributed across the entire tissue (figure 2). Then, to obtain compartment-specific protein measurements, we generated molecularly defined tissue compartments in each ROI by fluorescent marker colocalization. In YTMA471, three compartments were generated: tumor compartment (panCK+), macrophage compartment (CD68+), and lymphocyte compartment (panCK–/CD45+/CD68–). An additional fourth compartment containing the summation of the protein counts from lymphocyte and macrophage compartments (panCK–/CD45+/CD68+) was defined as the immune compartment in this cohort (figure 1). In H12O_ITX1 and CIMA-CUN cohorts, the tumor compartment (panCK+) and the immune compartment (panCK–/CD45+) were generated (figure 2). Subsequently, oligos from these compartments were released on ultraviolet light exposure, dispensed in a 96-well plate, hybridized to 4-color, 6-spot optical barcodes, and digitally counted in the nCounter system (NanoString). Using the GeoMx software (NanoString),

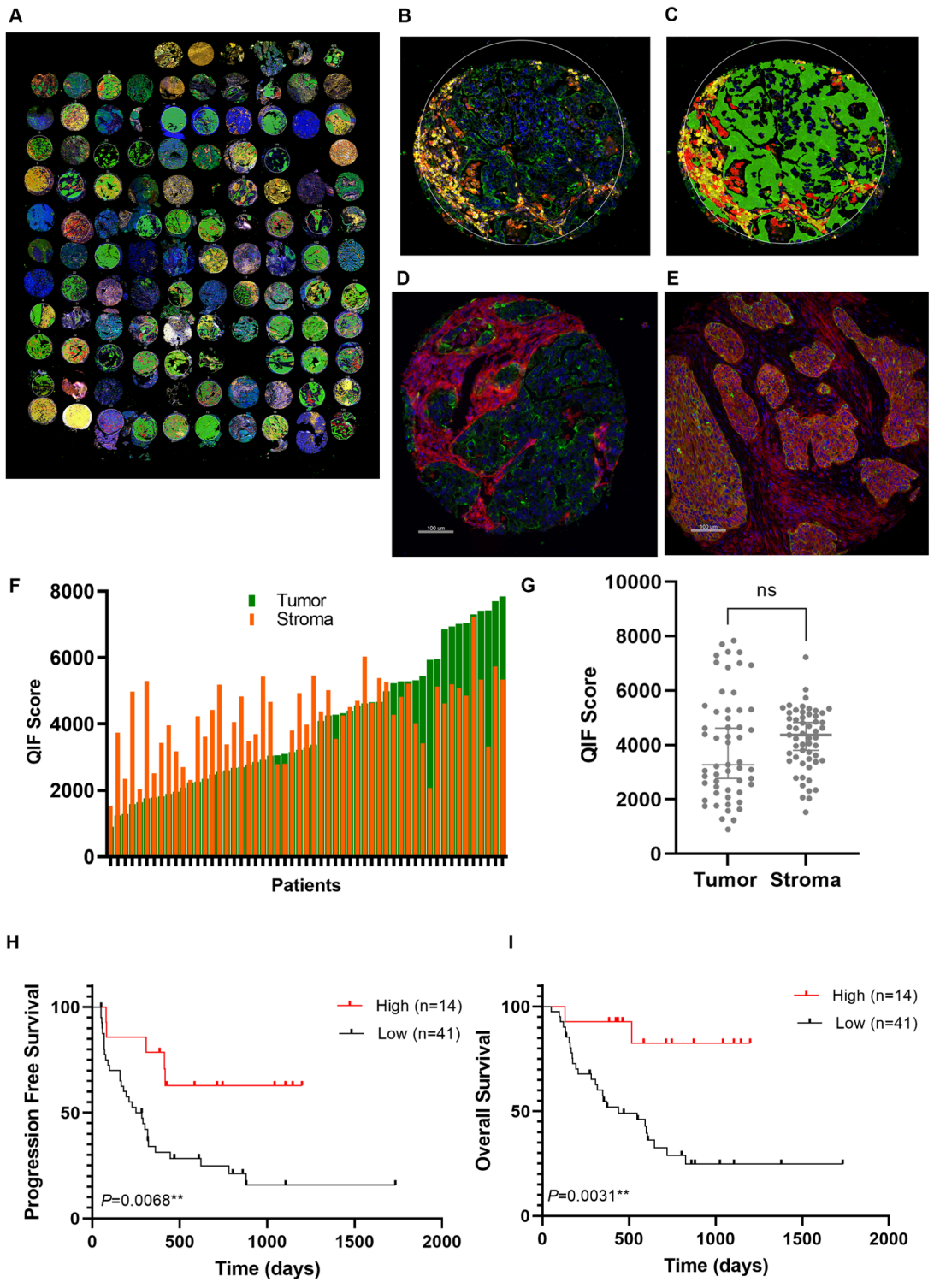


Figure 1 Identification of CD44 expression in the tumor compartment as a predictor of survival in YTMA471 discovery cohort. (A) Representative image of YTMA471 acquired using the GeoMx DSP system. (B–C) Representative TMA spot showing the fluorescence image (B) and the compartmentalized image created by fluorescence colocalization (C) using the GeoMx DSP system; panCK (green), CD45 (yellow), CD68 (red), SYTO13 (blue). (D–E) Representative TMA spot of low CD44 expression in panCK +tumor cells (D) and high CD44 expression in panCK +tumor cells (E) using QIF; panCK (green), CD44 (red), DAPI (blue). (F) Dynamic range of CD44 expression in the tumor compartment (panCK+) and in the stromal compartment (panCK–) using QIF. (G) Comparative analysis of CD44 levels measured by QIF in the tumor compartment and the stromal compartment. (H–I) Kaplan-Meier PFS curve (H) and OS curve (I) according to CD44 expression in the tumor compartment using QIF (optimal quartile cutpoint). CK, cytokeratin; DSP, digital spatial profiling; ns, not significant; OS, overall survival; PFS, progression-free survival; QIF, quantitative immunofluorescence; TMA, tissue microarray.

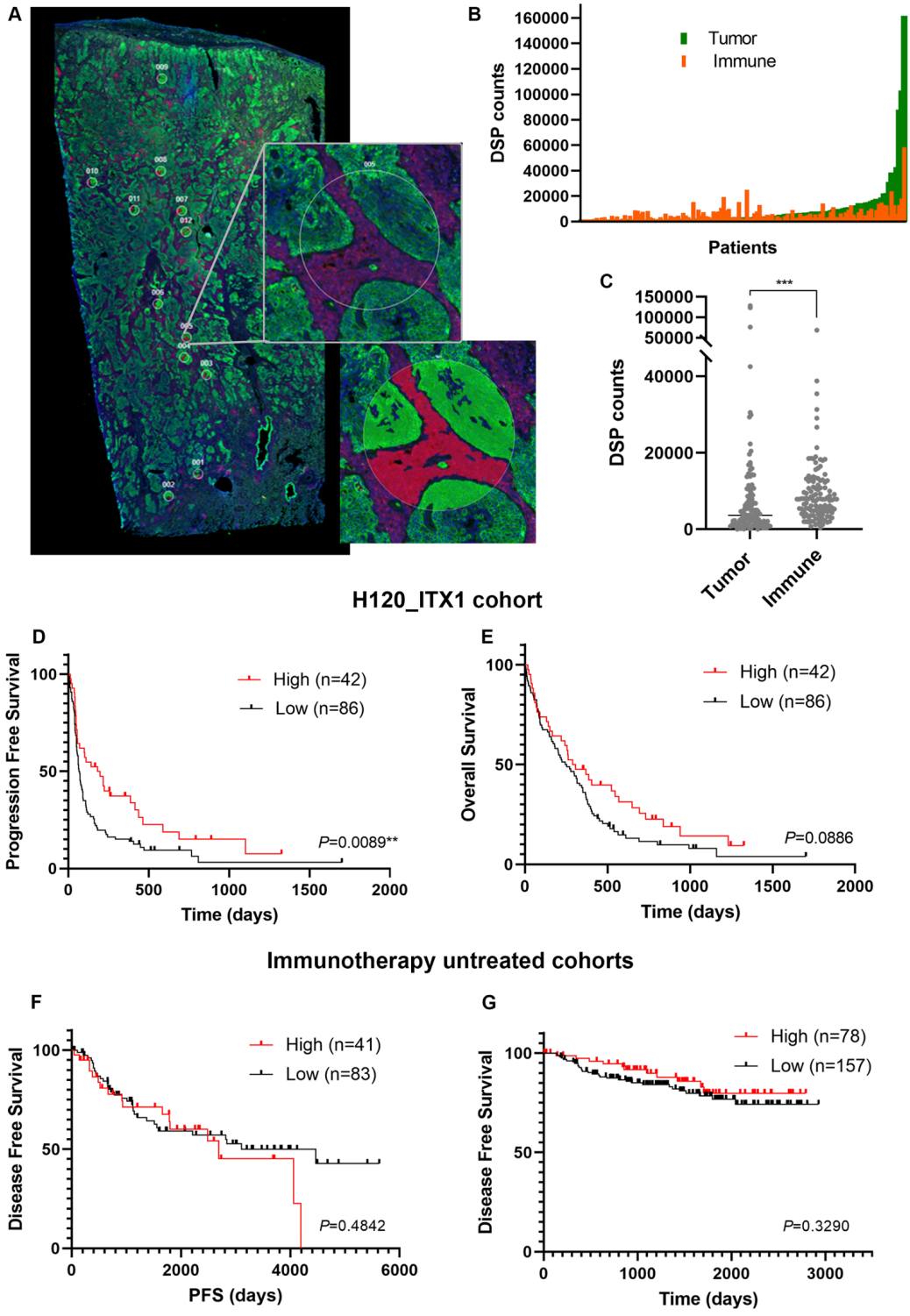


Figure 2 Validation of CD44 expression in the tumor compartment as an indicative biomarker of sensitivity to single-agent PD-1 axis blockade in NSCLC. (A) Representative image of a whole tissue section with 12 ROIs selected from immune-enriched (panCK+/CD45+) intratumoral areas using DSP. Fluorescence image is shown on the top, and the compartmentalized image at the bottom; panCK (green), CD45 (red), SYTO13 (blue). (B), Dynamic range of CD44 expression in the tumor compartment (panCK+) and in the immune compartment (panCK–/CD45+) using DSP; (C) Comparative analysis of CD44 levels measured by DSP in the tumor compartment and in the immune compartment. (D–E) Kaplan-Meier PFS curve (D) and OS curve (E) according to CD44 expression in the tumor compartment using DSP (tertile cutpoint). (F–G) Kaplan-Meier disease-free survival curves according to CD44 expression using DSP (F) or QIF (G) in immunotherapy untreated cohorts (CIMA-CUN cohort and YTMA423 cohort as F and G, respectively) (tertile cutpoint). CK, cytokeratin; DSP, digital spatial profiling; ns, not significant; NSCLC, non-small-cell lung cancer; OS, overall survival; PD-1, programmed cell death protein-1; PFS, progression-free survival; QIF, quantitative immunofluorescence; ROI, regions of interest.

digital counts were first normalized to internal spike-in controls (ERCCs), and then normalized to the counts of internal housekeeper protein probes. A more detailed description of the DSP protocol can be found in online supplemental materials and methods.

Quantitative immunofluorescence

We developed a quantitative immunofluorescence (QIF) protocol (CD44/CK/DAPI) to detect and quantify CD44 using slides from YTMA471 and YTMA423 cohorts. Two slides derived from independent TMA blocks, each block containing two non-adjacent tumor cores per NSCLC case, were used in both cohorts. The QIF staining protocol is detailed in online supplemental materials and methods.

We used the AQUA system (Navigate BP) to acquire and quantify the fluorescence signal of CD44 as previously described.[15] CD44 was measured from two compartments: the CK +tumor compartment, created by binarizing the CK signal, and the stromal compartment, created by excluding the tumor mask from a dilated DAPI mask representing the total tissue. CD44 QIF scores were calculated by dividing the target pixel intensity by the area of the compartment, and then normalized to the exposure time and bit depth at which the images were captured.[15]

CD44 antibody validation

To assess antibody specificity, we followed the guidelines for pillars of validation proposed by Uhlen *et al*.[16] We tested two independent anti-CD44 antibodies (clone 156–3 C11 (Cell Signaling) and clone EPR1013Y, (Abcam)) targeting non-overlapping epitopes (online supplemental table S3) in an NSCLC test array (YTMA295), which contained 35 lung tumor cores with variable CD44 expression. For each antibody, first we assessed membrane localization, then optimized the titration, and then compared the CD44 QIF scores obtained with the optimal antibody concentration of each clone. A high correlation in the quantitative measurements of CD44 between independent antibodies would prove specificity for recognizing the target.[16 17]

Immunotherapy efficacy assessment

We used Response Evaluation Criteria in Solid Tumors V.1.1 to retrospectively evaluate treatment response to immune checkpoint blockade. We defined clinical benefit as having experienced complete or partial response, or stable disease lasting ≥6 months, whereas non-clinical benefit was defined as primary progressive disease or stable disease lasting <6 months. Patients with stable disease who did not progress and were censored before 6 months of follow-up were non-evaluable. Overall survival (OS) and progression-free survival (PFS) were calculated from the treatment start date to the date of death or loss of follow-up, or the date of disease progression, death, or loss of follow-up, respectively. For those patients who did not die or progress during the study period, the outcome was considered left-censored.

Data processing and statistical analysis

To analyze the association between marker expression and patients' clinical-pathological characteristics or survival, normalized counts or QIF scores were averaged across all ROIs of each tissue sample to derive a single value per patient. We used non-parametric tests to compare CD44 counts between patient clinical subgroups or tissue compartments. For survival prediction analysis (PFS and OS), we used the Cox proportional hazards model, first utilizing averaged counts or QIF scores as continuous $\log_2$-transformed data. We used the X-tile software[18] to explore the optimal cutpoint for patients' survival stratification. Survival curves were computed with the Kaplan-Meier method and compared using the log-rank test. For prediction of clinical benefit (binary outcome) in the validation set, we utilized $\log_2$-transformed counts from each ROI independently (without averaging across ROIs) and fitted a regular logistic regression model with clustering by patient using Stata V.17.0 software. To test the association between CD44 counts in the tumor compartment with other immune-related proteins measured in the same compartment, we selected those ROIs containing matched tumor and immune compartment protein measurements and considered each ROI as an independent sample, since we aimed to assess the immune microenvironment features in each of the independently profiled intratumoral regions. Then, we used two complementary analyses. First, we used protein counts as continuous variables and tested the association between tumor cell CD44 with each of the other immune related markers using Pearson correlation analysis. Second, we dichotomized tumor cell CD44 expression using a clinically relevant cutpoint (upper tertile) to analyze what proteins were differentially expressed in tumor cell CD44 high versus tumor cell CD44 low ROIs. For this differential expression analysis, we quantified the fold change (FC) of all protein levels in ROIs with elevated CD44 expression relative to ROIs with low CD44 expression, and tested the significance with non-parametric tests accounting for multiple comparisons (Benjamini-Hochberg false discovery rate (FDR) method) using MetaboAnalyst V.5.0 software. All hypothesis testing was performed at a two-sided significance level of α=0.05.

RESULTS

We first assessed the YTMA471 discovery cohort using the DSP system. One-hundred and ten ROIs were generated in 56 NSCLC cases. Seventy-one proteins were measured selectively from four compartments (figure 1A–C), resulting in a total of 284 protein variables per ROI. Using continuous log-scaled data, PD-L1 expression in the tumor compartment (HR=0.67, p=0.017) and in the immune compartment (HR=0.52, p=0.008) was predictive of longer PFS, thus validating the cohort. Among the novel candidate predictors, CD44 expression measured in the tumor compartment, but not in stromal compartments, was associated with longer PFS both in the univariate

(HR=0.76, 95% CI 0.61 to 0.96; p=0.024) and in the multivariate analysis after controlling for stage, baseline liver metastasis, and Lung Immune Prognostic Index (LIPI) score[19] (HR=0.68, 95% CI 0.46 to 0.99; p=0.043) (online supplemental table S4), and was selected for further study and validation.

In the quantitative analysis using DSP, CD44 levels followed a continuous distribution. CD44 counts were significantly higher in the immune compartment than in the tumor compartment (Wilcoxon paired test, p<0.001) (online supplemental figure S1A). CD44 levels in the tumor compartment were significantly higher in patients with squamous-cell carcinomas (SCC) (Kruskal-Wallis test, p=0.0014) and patients with no baseline liver metastasis (Mann-Whitney test, p=0.0026). Tumor CD44 expression was not significantly different among other clinical factors analyzed (online supplemental figure S1B–H).

As the initial step for validation, we aimed to reproduce the DSP findings using an orthogonal QIF-based method in the same cohort. First, we sought to prove specificity for detecting CD44. For this purpose, we tested two different anti-CD44 antibodies that recognized non-overlapping epitopes in an NSCLC test array (YTMA295). We observed a good dynamic range (online supplemental figure S2A,B) and a specific membranous CD44 staining pattern with both antibodies (online supplemental figure S2C,D). Then, we compared the CD44 QIF scores obtained with the two clones, showing a high correlation coefficient (R^2=0.88) (online supplemental figure S2E). This provided evidence that both antibodies specifically recognized CD44 without significant cross reactivity, and we selected clone 156–3 C11 as a validated antibody for the remainder of the QIF studies. CD44 quantitative measurements were concordant between QIF and DSP methods (R^2=0.44) (online supplemental figure S2F).

Then, we evaluated CD44 expression patterns and outcome performance using QIF in YTMA471 cohort. CD44 displayed predominantly membranous staining pattern and was detected both in tumor cells and stromal cells (figure 1D,E). Virtually all tumors were visually positive for CD44 in the stroma, while 70% of NSCLCs had CD44 staining in tumor cells. CD44 QIF scores also followed a continuous distribution and were numerically higher in the stromal compartment than in the tumor compartment (Wilcoxon paired test, p=0.13) (figure 1F,G). Consistent with DSP findings, CD44 levels in the tumor compartment assessed as continuous QIF scores were significantly associated with longer PFS (multivariate HR=0.31, 95% CI 0.11 to 0.87; p=0.022) and OS (multivariate HR=0.29, 95% CI 0.09 to 0.97; p=0.038). Using the X-tile software, the highest quartile was determined as the optimal single cutpoint that significantly predicted both PFS (log-rank test, p=0.0068) and OS (log-rank test, p=0.0031) in this cohort (figure 1H,I). CD44 QIF levels in the stromal compartment did not predict outcomes, either as continuous data or using median, tertile, or quartile cutpoints (online supplemental figure S3A–F).

Next, after having validated tumor cell CD44 as a candidate predictor using two orthogonal methods in the discovery set, we sought for external validation in the H12O_ITX1 cohort using DSP. We generated a total of 662 ROIs from 128 patient-derived whole tissue sections (average ROI per tissue sample=6 (1–12)), each ROI was then segmented in the tumor compartment and the immune compartment (figure 2A). In this validation set, PD-L1 as well as other known markers such as CD3 or CD8 were associated with outcomes from PD-1 axis blockade (online supplemental figure S4A–F), thus providing validity to the cohort and the ROI selection approach.

Consistent with the results in the YTMA471 discovery cohort, CD44 expression in the tumor compartment followed a continuous distribution (figure 2B) and was significantly higher in the immune compartment than in the tumor compartment (Wilcoxon paired test, p<0.001) (figure 2C). As seen in the training set, tumor cell CD44 expression was significantly higher in patients with SCC (Kruskal-Wallis test, p<0.001) (online supplemental figure S5A–F). In terms of outcome, CD44 expression in the tumor compartment validated as a predictor of clinical outcomes in the H12O_ITX1 cohort. Assessed as a continuous variable, CD44 levels in the tumor compartment were predictive of clinical benefit (OR=1.22, 95% CI 1.03 to 1.45; p=0.018). Using the top tertile cutpoint, patients with high CD44 expression in the tumor compartment showed significantly longer PFS under single-agent PD-1 axis blockade (HR=0.58, 95% CI 0.38 to 0.87; p=0.01) (figure 2D). In the multivariate analysis, CD44 expression in the tumor compartment remained as an independent predictor of PFS after adjusting for performance status, baseline liver metastasis, and LIPI score (HR=0.62, 95% CI 0.40 to 0.96; p=0.035). The effect of tumor cell CD44 in predicting PFS remained significant after correcting for PD-L1 Tumor Proportion Score (TPS), either using the PD-L1 ≥1% TPS cutpoint (HR=0.25, p=0.003 in the discovery set; HR=0.59, p=0.015 in the validation set), or the PD-L1 ≥50% TPS cutpoint (HR=0.26, p=0.004 in the discovery set; HR=0.62, p=0.032 in the validation set) (online supplemental table S5). In this regard, when stratifying patients with NSCLC by clinically relevant cutpoints for tumor cell CD44 and PD-L1 TPS, there were approximately 45% of discordant cases when using the PD-L1 ≥1% TPS cutpoint, and about 30% of discordant cases when using the PD-L1 ≥50% TPS cutpoint consistently in both cohorts (online supplemental figure S6A–D). In the validation set, the differences in OS according to tumor cell CD44 levels were not statistically significant (log-rank test, p=0.08) (figure 2E). CD44 expression in the immune compartment showed no association with clinical outcomes (online supplemental figure S7A,B).

Then, to assess the prognostic value of tumor cell CD44 expression, we analyzed two control cohorts of patients with NSCLC that never received PD-1 axis inhibitors during the whole course of their disease. The quantitative analysis of CD44 expression in these

cohorts confirmed significantly higher CD44 levels in the immune compartment, and higher CD44 expression in the tumor compartment in patients with SCC (online supplemental figure S8A–D). In contrast to the immunotherapy treated cohorts, CD44 expression in the tumor compartment, assessed either by DSP (CIMA-CUN), or by QIF (YTMA423), was not predictive of survival outcomes in the absence of PD-1 axis blockade (figure 2F,G).

Finally, we explored whether a distinct tumor microenvironment could potentially explain a higher responsiveness to PD-1 axis blockade of NSCLC cases with high tumor cell CD44 expression. For this analysis, we used ROI-level counts in the tumor compartment from two cohorts, H12O_ITX1 and CIMA-CUN, as they were assessed with the same DSP assay and covered a greater number of ROIs in each tissue sample. First, we tested the correlation between tumor cell CD44 expression and each individual marker as continuous variables. Among all the markers with statistically significant results (online supplemental table S6), CD80 showed the highest, although modest, positive correlation (R^2 >0.2) consistently in both cohorts (figure 3A,B). There was a statistically significant (p<0.0001) but modest correlation (R^2 <0.3) between tumor cell CD44 counts and PD-L1 counts in the CIMA-CUN cohort, but this was not reproduced in the H12O_ITX1 cohort (figure 3C,D). In the H12O_ITX1 data set, the association between tumor cell CD44 and PD-L1 was significant (p=0.005) when removing extreme CD44 and PD-L1 values falling outside the upper limits of expression in the CIMA-CUN data set, but still with a loose correlation coefficient (R^2=0.01) (online supplemental figure S9). Then, we interrogated what proteins were differentially expressed in the ROIs with high CD44 expression in the tumor compartment (top tertile) as compared with those with low CD44 expression (rest). Consistently in the two cohorts, CD44 high ROIs showed prominent and statistically significant upregulation (FC >1.5, FDR-adjusted p<0.05) of PD-L1, TIM-3, ICOS, and CD40. Other immune cell markers (CD45, CD8, CD11c), co-inhibitory molecules (VISTA, B7-H3), and co-stimulatory molecules (CD27), were also enriched in ROIs with elevated CD44 expression in both cohorts, but were less prominently upregulated (figure 3E,F and online supplemental table S7).

DISCUSSION

In the present study, we have discovered and validated tumor cell CD44 as a new biomarker associated with sensitivity to PD-1 axis blockade across multiple NSCLC cohorts. Using two quantitative and spatially informed technologies in independent NSCLC cohorts, we show that CD44 overexpression in the tumor compartment, but not in the immune compartment, predicts clinical outcomes from PD-1 axis blockade. By analyzing two additional immunotherapy untreated cohorts, we found that the association between tumor cell CD44 and outcome was specific for anti-PD-1/PD-L1 treatment. Finally, we show that NSCLCs with CD44 overexpression in the tumor compartment are associated with a unique immune microenvironment.

CD44 plays multifaceted roles in cancer biology, including modulation of antitumor immunity. It is a surface marker commonly overexpressed by tumor cells with cancer initiating properties,[6 7 20] and has therefore been repeatedly associated with cancer stem cells particularly in breast cancer.[20–24] To our knowledge, the association between tumor cell CD44 overexpression and improved clinical outcomes to PD-1 axis blockade has not been previously reported. In the present study, we have performed rigorous validation of this novel finding, including orthogonal biomarker assessment with a second QIF-based technology (with prior validation of an anti-CD44 antibody following recommended guidelines[16]), and external validation utilizing whole tissue sections by independent investigators. It is possible that the use of a technology that provides quantitative data considering the cell compartment of expression might have increased the sensitivity to detect predictive information in our study. For instance, there are several studies that have evaluated the association between bulk tumor transcriptomes and response to PD-1 axis blockade, where CD44 is not among the list of overexpressed genes linked to response and PFS.[25] However, as CD44 protein is quantitatively more abundant in the immune compartment than in the epithelial compartment, the predictive capacity of tumor cell CD44 might have been diluted when using bulk mRNA analysis[25] or more subjective and semiquantitative methods such as IHC. Using spatially resolved quantitative protein measurements, we also excluded prognostic relevance of tumor cell CD44 expression. Although some studies have reported that CD44 overexpression is associated with poor prognosis (particularly overexpression of variant isoform 6 (CD44v6)) in some tumor types,[26 27] these studies have not used quantitative assays enabling cell compartment specific protein measurements.

The higher sensitivity to PD-1 axis inhibition observed in NSCLCs with CD44 overexpression by cancer cells in this study could be at least partially explained by the role of CD44 in regulating antitumor immunity. CD44 has been identified as a PD-L1 modulator in preclinical models.[13 14] Notably, CD44 expressing cancer cells derived from human squamous cell carcinoma of the head and neck showed higher responsiveness to IFN-γ than CD44 negative cells, which resulted in greater inducible PD-L1 expression following IFN-γ exposure in vitro.[13] Consistent with these preclinical findings, we did observe a statistically significant association between tumor cell CD44 and PD-L1, but the strength of this association, particularly the linear relationship between both markers when assessed as continuous values, is low-to-modest at best. From a clinical perspective, this means that there is non-redundancy or at least some degree of complementarity between both biomarkers for patient stratification. Beyond its potential role in PD-L1 modulation, and further supporting a role for CD44 in adaptive immune resistance, it has been

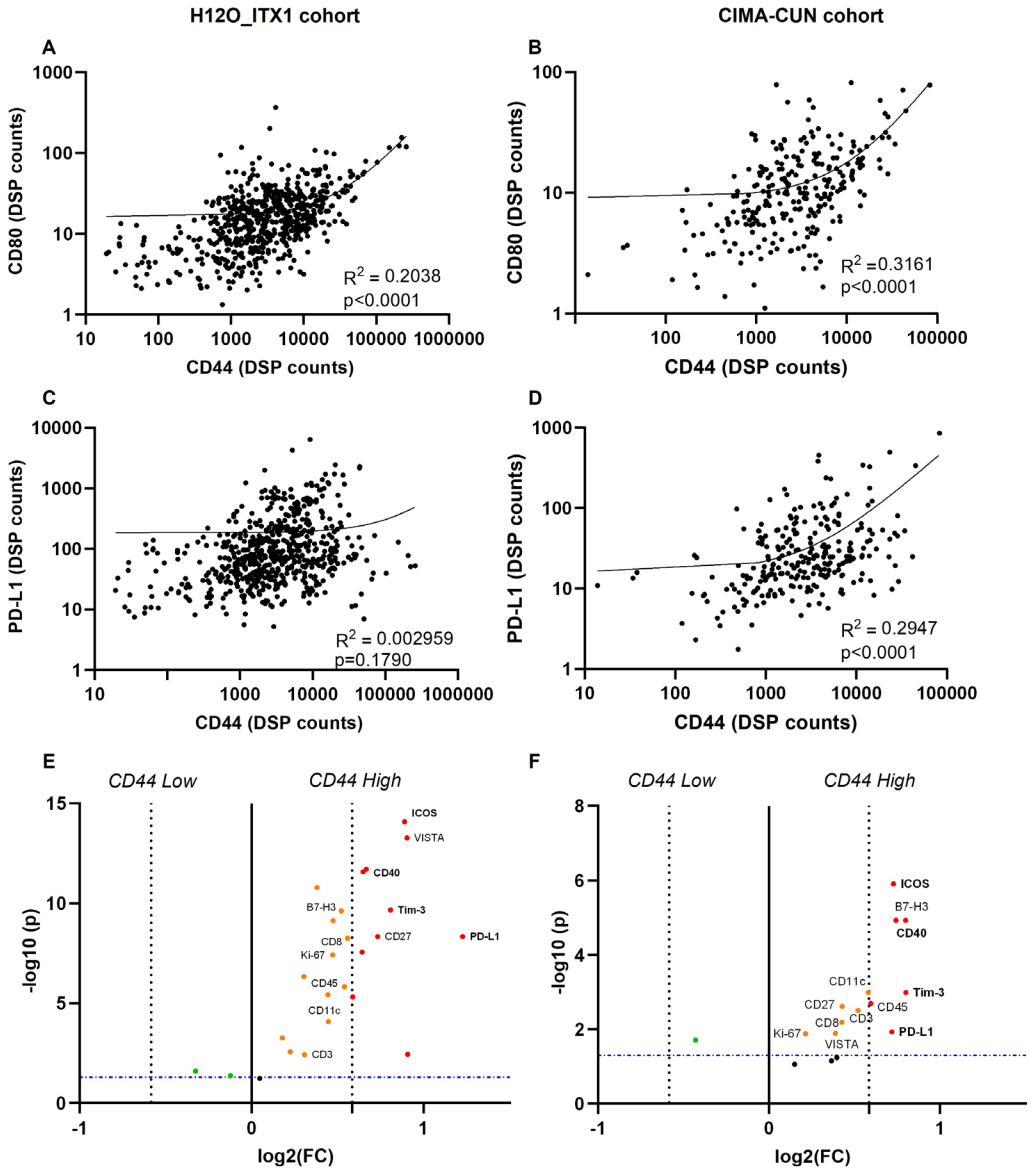


Figure 3 CD44 levels in the tumor compartment and immune microenvironment features in human NSCLC. (A–B) Correlation between CD44 and CD80 levels in the tumor compartment in H12O_ITX1 cohort (A) and CIMA-CUN cohort (B). (C–D) Correlation between CD44 and PD-L1 levels in the tumor compartment in H12O_ITX1 cohort (C) and CIMA-CUN cohort (D). (E–F) Differentially expressed protein markers in ROIs with high CD44 expression in the tumor compartment (top tertile) relative to ROIs with low CD44 expression in the tumor compartment (rest) in H12O_ITX1 cohort (E) and CIMA-CUN cohort (F). The significance (FDR-adjusted p values) is represented relative to the FC in protein levels in CD44 high relative to CD44 low ROIs. Only statistically significant markers are highlighted. Markers with a FC >1.5 and FDR-adjusted p values<0.05 in the two cohorts are marked in red bold. DSP, digital spatial profiling; FC, fold change; FDR, false discovery rate; PD-L1, programmed death-ligand 1; NSCLC, non-small-cell lung cancer; ROI, regions of interest.

shown that tumor initiating cancer cells expressing CD44 can selectively upregulate CD80 under immune pressure, thereby suppressing T cell immunity in a CTLA4-dependent manner.[28] This is also concordant with the significant correlation observed between CD44 and CD80 counts in the tumor compartment in our study. On the other hand, CD44 is a necessary protein for EMT,[5–9] and EMT has been associated with a broad inflammatory microenvironment characterized by coexistent upregulation of co-stimulatory and co-inhibitory checkpoint molecules in human lung cancer.[24 29–31] We also observed that intratumoral regions with elevated tumor cell CD44 had prominent upregulation of co-inhibitory (TIM-3, B7-H3, PD-L1) and co-stimulatory (ICOS, CD40, CD27) molecules in two independent NSCLC cohorts. Of note, although EMT has been generally associated with poorer outcomes and treatment resistance in several cancers,[6 7] some tumor types with more mesenchymal features such as sarcomatoid malignant pleural mesotheliomas[32] or certain subtypes of small cell lung cancers (SCLC),[33] also

display an inflammatory microenvironment and show higher sensitivity to checkpoint blockade therapies.[33 34] Whether tumor cell CD44 plays a role in determining sensitivity to immunotherapy in these tumor types remains unknown, but might deserve to be explored in future studies. Collectively, these data suggest that NSCLCs with CD44 overexpression in cancer cells contain a distinct tumor microenvironment that is potentially more prone to immune checkpoint blockade responsiveness.

From a clinical standpoint, it needs to be considered that with the current use of chemoimmunotherapy in the front-line setting of advanced NSCLC, where even PD-L1 has no or limited utility to select best candidates to receive PD-1 axis inhibitors, it is unlikely that tumor cell CD44 will have an immediate major clinical impact as a biomarker for patient stratification. However, as quantitative tumor cell CD44 expression appears non-redundant with PD-L1 expression, there is potential to use them as complementary biomarkers. This possibility seems particularly appealing in tumor types where PD-L1 offers limited predictive information.[34–37] On the other hand, CD44 is a surface glycoprotein potentially amenable for therapeutic antibody-based blockade. Some attempts to target CD44 using monoclonal antibodies have shown little clinical activity,[27 38] but based on our findings, CD44 could be reconsidered as a target for tumor-cell-specific antibody-based strategies particularly in combination with PD-L1 blockade (ie, bi-specific antibodies with *cis* co-engagement[39] or anti-CD44 probody formats[40]), and preferentially in those cases with specifically tumor cell CD44 overexpression.

This study needs to be interpreted in the context of a number of limitations. First, we analyzed tumor samples collected from retrospective immunotherapy treated cohorts, not randomized clinical trials. Retrospective studies have limitations including the possibility of selection bias, so despite having found tumor cell CD44 overexpression as a marker associated with sensitivity to PD-1 axis blockade in two independent cohorts, these results still need to be interpreted with caution and can only be considered hypothesis generating. In this same line, in the absence of randomization, we could not perform an interaction test between CD44 and PD-1 axis inhibitor treatment in a statistical model to prove that CD44 is a predictive biomarker. We were able to show absence of prognostic value and thus have used the term 'indicative' rather than predictive. Second, the tumor samples used for this study were archived samples from different tumor sites, not collected immediately preceding immunotherapy in most of the cases. The optimal time to assess immunotherapy biomarkers is still undefined, and some studies suggest that the key biological information is reasonably preserved in archived tissue.[41] Another limitation is that we do not fully validate an identical cutpoint in the training and validation sets for optimal patient stratification. Future studies addressing the predictive value of CD44 will require the validation of an optimal and reproducible cutpoint in larger well-powered cohorts. It is perhaps also a limitation that we could not correct the CD44 survival prediction analysis by TMB since most of the patients in our cohorts did not have their tumors assessed for TMB status. Finally, this study does not prove the mechanism by which NSCLCs with CD44 overexpression in the tumor compartment are more sensitive to PD-1 axis blockade. Although the unique immune microenvironment of these tumors might partially explain a higher responsiveness, further functional studies are needed to understand the interplay between CD44 positive cancer stem cell phenotype and mechanisms of immune evasion.

In conclusion, this works highlights a previously underappreciated role for tumor cell CD44 and sensitivity to PD-1 axis blockade in NSCLC. With further validation, CD44 could complement existing biomarkers for optimal patient stratification, and potentially open new therapeutic strategies to improve precision immunotherapy in lung cancer.

Author affiliations
[1]Department of Pathology, Yale University School of Medicine, New Haven, Connecticut, USA
[2]Tumor Microenvironment and Immunotherapy Research Group, 12 de Octubre Research Institute (i+12), Madrid, Spain
[3]H12O-CNIO Lung Cancer Clinical Research Unit, Instituto de Investigación Sanitaria Hospital 12 de Octubre (imas12), Spanish National Cancer Research Center (CNIO), Madrid, Spain
[4]Department of Medical Oncology, 12 de Octubre Hospital, Madrid, Spain
[5]Program in Solid Tumors, CIMA-University of Navarra, Pamplona, Spain
[6]Department of Pathology, University of Navarra, Pamplona, Spain
[7]Spanish Center for Biomedical Research Network in Oncology, CIBERONC, Madrid, Spain
[8]Health Research Institute of Navarra, IdiSNA, Pamplona, Spain
[9]Department of Medicine (Oncology), Yale University School of Medicine, New Haven, Connecticut, USA
[10]Department of Medicine, Complutense University, Madrid, Spain

Twitter Myrto K Moutafi @mkmoutafi

Acknowledgements The authors thank Lori A Charette and the staff of Yale Pathology tissue services for expert histology services, and Teresa Muñoz for her help in reviewing tumor specimens for adequacy for Digital Spatial Profiling analysis.

Contributors MKM, MM, DLR, and JZ conceived the study. MKM, MM, JB, VA, LC-L performed experiments and/or data analysis. Clinical information and specimens were provided by MKM, JB, VA, DO, KV, AC, LMM, SPA, LP-A, DLR, and JZ. Manuscript preparation was performed by MKM, MM, DLR, and JZ. Manuscript revisions were performed by all authors. All authors approved the final manuscript. JZ is responsible for the overall content of the manuscript as the guarantor.

Funding This work was supported by grants from the Lung Cancer Foundation of America (LCFA/IASLC/BMS research award), Spanish Association Against Cancer (AECC, LABAE2044), and Carlos III Health Institute (ISCIII, PI20/01494) to JZ, and Yale SPORE in Lung Cancer (P50 CA196530) to DLR and RSH. MKM is receiving a scholarship from the Hellenic Society of Medical Oncologists (HESMO). LMM is supported by FIMA, AECC Scientific Foundation (GCB14-2170), Fundación Ramón Areces, Spanish Ministry of Economy and Innovation, Fondo de Investigación Sanitaria-Fondo Europeo de Desarrollo Regional (PI19/00098), and Fundación Roberto Arnal Planelles. LP-A is supported by CIBERONC (CB16/12/00442). DLR is supported by the Yale SPORE in Lung Cancer (P50 CA196530) and has a sponsored research agreement from Konica/Minolta. JZ is supported by a Juan Rodes contract from the Carlos III Health Institute (JR19/00028).

Competing interests SMM is currently a Boehringer Ingelheim employee. LMM reports speakers honoraria from AstraZeneca and research grants from AstraZeneca and BMS. KAS reports consulting or advisory roles for Shattuck Labs, Pierre Fabre, EMD Serono, Clinica Alemana de Santiago, Genmab, Takeda, Merck Sharpe & Dohme, Bristol Myers Squibb, AstraZeneca, Agenus, Repertoire

Therapeutics, OnCusp and Ariagen. Reports grants or research funding from Navigate Biopharma, Tesaro/GSK, Moderna, Takeda, Surface Oncology, Pierre Fabre Research Institute, Merck Sharpe & Dohme, Bristol Myers Squibb, AstraZeneca, Ribon Therapeutics, Akoya Biosciences, Boehringer Ingelheim and Eli Lilly. RSH has served as Non-Executive Director for Immunocore and is a member of the Board of Directors (non-executive/independent) for Junshi Biosciences; is a consultant for AbbVie, Armo Biosciences, AstraZeneca, Bristol Myers Squibb, Bayer HealthCare Pharmaceuticals, Bolt Biotherapeutics, Candel Therapeutics, Checkpoint Therapeutics, Cybrexa Therapeutics, DynamiCure Biotechnology, Eli Lilly and Company, eFFECTOR Therapeutics, EMD Serono, Foundation Medicine, Genentech/Roche, Genmab, Gilead, HiberCell, I-Mab Biopharma, Immune-Onc Therapeutics, Immunocore, Johnson & Johnson, Loxo Oncology, Merck and Company, Mirati Therapeutics, NextCure, Novartis, Ocean Biomedical, Oncocyte, Oncternal Therapeutics, Pfizer, Refactor Health, Ribbon Therapeutics, Sanofi, STCube Pharmaceuticals, Takeda, WindMIL Therapeutics, Xencor; has received research support from AstraZeneca, Eli Lilly and Company, Genentech/Roche, and Merck and Company; is a committee chair in American Association for Cancer Research, International Association for the Study of Lung Cancer, Society for Immunotherapy of Cancer, Southwest Oncology Group LP-A reports receiving honoraria from Amgen, AstraZeneca, Bayer, Blueprint Medicines, Bristol Myers Squibb, Celgene, Ipsen, Eli Lilly, Merck Serono, Mirati Therapeutics, Merck Sharp & Dohme, Novartis, Pfizer, PharmaMar, Roche/Genentech, Sanofi, Servier, and Takeda; leadership fees from Genomica and ALTUM Sequencing; research funding from AstraZeneca, Bristol Myers Squibb, Kura Oncology, PharmaMar, and Merck Sharp & Dohme; speaker fees from Bristol Myers Squibb, Eli Lilly, Merck Serono, Merck Sharp & Dohme Oncology, Pfizer, and Roche/Genentech; and travel, accommodation, and expenses from AstraZeneca, Bristol Myers Squibb, Merck Sharp & Dohme, Pfizer, Roche, and Takeda. DLR reports grants from Navigate Biopharma and Konica/Minolta/Invicro during the conduct of the study. Activities outside this work include honoraria and/or grants and/or instrument support from Akoya, Amgen, AstraZeneca, BMS, Cell Signaling Technology, Cepheid, Danaher, Konica/Minolta, Lilly, Merck, NanoString, NextCure, Odonate, Paige.AI, Roche, Sanofi, and Ventana. JZ has served as a consultant for AstraZeneca, BMS, Roche, Pfizer, Novartis, and Guardant Health. Reports speakers honoraria from BMS, Pfizer, Roche, AstraZeneca, NanoString and Guardant Health. Reports travel honoraria from BMS, Pfizer, Roche, AstraZeneca, and NanoString. Receives research support/funds from BMS, AstraZeneca, and Roche. The rest of the authors declare no conflicts of interest.

Patient consent for publication Not applicable.

Ethics approval All study procedures were done under the approval of the Yale Human Investigation Committee protocol #9505008219, 12 de Octubre Ethics Committee protocols #18/062 and #20/153, and University of Navarra Research Ethics Committee protocol #19/150. Investigation committees approved patients' informed consent or in some cases waiver of consent all in accordance with ethical guidelines.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplemental information. Data are available upon reasonable request.

ORCID iDs
Myrto K Moutafi http://orcid.org/0000-0002-9314-3187
Ioannis A Vathiotis http://orcid.org/0000-0002-1772-5986
Niki Gavrielatou http://orcid.org/0000-0003-1380-6831
Gorka Ruiz de Garibay http://orcid.org/0000-0001-9936-8419
Alfonso Calvo http://orcid.org/0000-0003-4074-4242
Kurt A Schalper http://orcid.org/0000-0001-5692-4833
David L Rimm http://orcid.org/0000-0001-5820-4397
Jon Zugazagoitia http://orcid.org/0000-0001-9352-5882

REFERENCES

1 Doroshow DB, Sanmamed MF, Hastings K, *et al*. Immunotherapy in non-small cell lung cancer: Facts and hopes. *Clin Cancer Res* 2019;25:4592–602.
2 Grant MJ, Herbst RS, Goldberg SB. Selecting the optimal immunotherapy regimen in driver-negative metastatic NSCLC. *Nat Rev Clin Oncol* 2021;18:625–44.
3 Morad G, Helmink BA, Sharma P, *et al*. Hallmarks of response, resistance, and toxicity to immune checkpoint blockade. *Cell* 2021;184:5309–37.
4 Lu S, Stein JE, Rimm DL, *et al*. Comparison of biomarker modalities for predicting response to PD-1/PD-L1 checkpoint blockade: a systematic review and meta-analysis. *JAMA Oncol* 2019;5:1195–204.
5 Zöller M. CD44: can a cancer-initiating cell profit from an abundantly expressed molecule? *Nat Rev Cancer* 2011;11:254–67.
6 Lambert AW, Weinberg RA. Linking EMT programmes to normal and neoplastic epithelial stem cells. *Nat Rev Cancer* 2021;21:325–38.
7 Dongre A, Weinberg RA. New insights into the mechanisms of epithelial–mesenchymal transition and implications for cancer. *Nat Rev Mol Cell Biol* 2019;20:69–84.
8 Brown RL, Reinke LM, Damerow MS, *et al*. CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression. *J Clin Invest* 2011;121:1064–74.
9 Larsen JE, Nathan V, Osborne JK, *et al*. ZEB1 drives epithelial-to-mesenchymal transition in lung cancer. *J Clin Invest* 2016;126:3219–35.
10 Natsuizaka M, Whelan KA, Kagawa S, *et al*. Interplay between Notch1 and Notch3 promotes EMT and tumor initiation in squamous cell carcinoma. *Nat Commun* 2017;8:1758.
11 Wang Y-Y, Vadhan A, Chen P-H, *et al*. Cd44 promotes lung cancer cell metastasis through erk–zeb1 signaling. *Cancers* 2021;13:4057–15.
12 Guo J-Y, Hsu H-S, Tyan S-W, *et al*. Serglycin in tumor microenvironment promotes non-small cell lung cancer aggressiveness in a CD44-dependent manner. *Oncogene* 2017;36:2457–71.
13 Lee Y, Shin JH, Longmire M, *et al*. Cd44+ cells in head and neck squamous cell carcinoma suppress T-cell-mediated immunity by selective constitutive and inducible expression of PD-L1. *Clin Cancer Res* 2016;22:3571–81.
14 Kong T, Ahn R, Yang K, *et al*. CD44 promotes PD-L1 expression and its tumor-intrinsic function in breast and lung cancers. *Cancer Res* 2020;80:444–57.
15 Camp RL, Chung GG, Rimm DL. Automated subcellular localization and quantification of protein expression in tissue microarrays. *Nat Med* 2002;8:1323–8.
16 Uhlen M, Bandrowski A, Carr S, *et al*. A proposal for validation of antibodies. *Nat Methods* 2016;13:823–7.
17 MacNeil T, Vathiotis IA, Martinez-Morilla S, *et al*. Antibody validation for protein expression on tissue slides: a protocol for immunohistochemistry. *Biotechniques* 2020;69:460–8.
18 Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res* 2004;10:7252–9.
19 Mezquita L, Auclin E, Ferrara R, *et al*. Association of the lung immune prognostic index with immune checkpoint inhibitor outcomes in patients with advanced non-small cell lung cancer. *JAMA Oncol* 2018;4:351–7.
20 Creighton CJ, Li X, Landis M, *et al*. Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. *Proc Natl Acad Sci U S A* 2009;106:13820–5.
21 Al-Hajj M, Wicha MS, Benito-Hernandez A, *et al*. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* 2003;100:3983–8.
22 Prince ME, Kaczorowski RSA, Wolf GT. Identification of a subpopulation of cells with cancer stem cell properties. *Proc Natl Acad Sci U S A* 2007;104:973–8.
23 Nishino M, Ozaki M, Hegab AE, *et al*. Variant CD44 expression is enriching for a cell population with cancer stem cell-like characteristics in human lung adenocarcinoma. *J Cancer* 2017;8:1774–85.
24 Kursunel MA, Taskiran EZ, Tavukcuoglu E, *et al*. Small cell lung cancer stem cells display mesenchymal properties and exploit

immune checkpoint pathways in activated cytotoxic T lymphocytes. *Cancer Immunol Immunother* 2022;71:445–59.
25 Prat A, Navarro A, Paré L, *et al*. Immune-related gene expression profiling after PD-1 blockade in non–small cell lung carcinoma, head and neck squamous cell carcinoma, and melanoma. *Cancer Res* 2017;77:3540–50.
26 Zhao S, He J-L, Qiu Z-X, *et al*. Prognostic value of CD44 variant exon 6 expression in non-small cell lung cancer: a meta-analysis. *Asian Pacific J Cancer Prev* 2014;15:6761–6.
27 Chen C, Zhao S, Karnad A, *et al*. The biology and role of CD44 in cancer progression: therapeutic implications. *J Hematol Oncol* 2018;11:1–23.
28 Miao Y, Yang H, Levorse J, *et al*. Adaptive immune resistance emerges from tumor-initiating stem cells. *Cell* 2019;177:1172–86.
29 Chen L, Gibbons DL, Goswami S, *et al*. Metastasis is regulated via microRNA-200/ZEB1 axis control of tumour cell PD-L1 expression and intratumoral immunosuppression. *Nat Commun* 2014;5:5241.
30 Lou Y, Diao L, Cuentas ERP, *et al*. Epithelial-mesenchymal transition is associated with a distinct tumor microenvironment including elevation of inflammatory signals and multiple immune checkpoints in lung adenocarcinoma. *Clinical Cancer Research* 2016;22:3630–42.
31 Kim S, Koh J, Kim M-Y, *et al*. PD-L1 expression is associated with epithelial-to-mesenchymal transition in adenocarcinoma of the lung. *Hum Pathol* 2016;58:7–14.
32 Hmeljak J, Sanchez-Vega F, Hoadley KA, *et al*. Integrative molecular characterization of malignant pleural mesothelioma. *Cancer Discov* 2018;8:1548–65.
33 Gay CM, Stewart CA, Park EM, *et al*. Patterns of transcription factor programs and immune pathway activation define four major subtypes of SCLC with distinct therapeutic vulnerabilities. *Cancer Cell* 2021;39:346–60.
34 Baas P, Scherpereel A, Nowak AK, *et al*. First-line nivolumab plus ipilimumab in unresectable malignant pleural mesothelioma (CheckMate 743): a multicentre, randomised, open-label, phase 3 trial. *Lancet* 2021;397:375–86.
35 Goldman JW, Dvorkin M, Chen Y, *et al*. Durvalumab, with or without tremelimumab, plus platinum–etoposide versus platinum–etoposide alone in first-line treatment of extensive-stage small-cell lung cancer (CASPIAN): updated results from a randomised, controlled, open-label, phase 3 trial. *Lancet Oncol* 2021;22:51–65.
36 Liu SV, Reck M, Mansfield AS, *et al*. Updated overall survival and PD-L1 subgroup analysis of patients with extensive-stage small-cell lung cancer treated with Atezolizumab, carboplatin, and etoposide (IMpower133). *J Clin Oncol* 2021;39:619–30.
37 Fennell DA, Ewings S, Ottensmeier C, *et al*. Nivolumab versus placebo in patients with relapsed malignant mesothelioma (CONFIRM): a multicentre, double-blind, randomised, phase 3 trial. *Lancet Oncol* 2021;22:1530–40.
38 Menke-van der Houven van Oordt CW, Gomez-Roca C, van Herpen C, *et al*. First-in-human phase I clinical trial of RG7356, an anti-CD44 humanized antibody, in patients with advanced, CD44-expressing solid tumors. *Oncotarget* 2016;7:80046–58.
39 Blanco B, Domínguez-Alonso C, Alvarez-Vallina L. Bispecific immunomodulatory antibodies for cancer immunotherapy. *Cli Cancer Res* 2021;27:5457–64.
40 Autio KA, Boni V, Humphrey RW, *et al*. Probody therapeutics: an emerging class of therapies designed to enhance on-target effects with reduced off-tumor toxicity for use in immuno-oncology. *Clin Cancer Res* 2020;26:984–9.
41 Herbst RS, Baas P, Perez-Gracia JL, *et al*. Use of archival versus newly collected tumor samples for assessing PD-L1 expression and overall survival: an updated analysis of KEYNOTE-010 trial. *Ann Oncol* 2019;30:281–9.

Supplementary table S1. Baseline characteristics of the immunotherapy untreated NSCLC cohorts

	YTMA423	CIMA-CUN
Characteristic	**N (%)**	**N (%)**
Total	252	124
Gender	98 (39)	98 (79)
Male	154 (61)	26 (21)
Female		
Age		
<70 yo	141 (55.9)	88 (71)
≥70 yo	111 (44.1)	36 (29)
Smoking history		
Never smoker	31 (12)	15 (12)
Current smoker	62 (25)	28 (22.7)
Former smoker	159 (63)	81 (65.3)
Histology		
Adenocarcinoma	174 (69)	67 (54)
Squamous-cell carcinoma	66 (26)	46 (37.1)
Large-cell carcinoma	6 (2.4)	5 (4)
Other	5 (2)	6 (4.9)
Stage		
I	200 (79)	55 (44.4)
II	37 (14.7)	30 (24.2)
III	12 (4.8)	34 (27.4)
IV	3 (1.2)	5 (4)

Supplementary Table S2. Human IO DSP panels for protein detection used in this study

Core/Module	Target	YTMA471 cohort	H12O_ITX1 cohort	CIMA-CUN cohort
Immune Cell Profiling Panel				
Human Protein Core	Beta-2-microglobulin	X	X	X
	CD11c	X	X	X
	CD20	X	X	X
	CD3	X	X	X
	CD4	X	X	X
	CD45	X	X	X
	CD56	X	X	X
	CD68	X	X	X
	CD8	X	X	X
	CTLA4	X	X	X
	Granzyme B (GZMB)	X	X	X
	Ki-67	X	X	X
	PD-1	X	X	X
	PD-L1	X	X	X
	Pan-cytokeratin	X	X	X
	HLA-DR	X	X	X
	SMA	X	X	X
	Fibronectin	X	X	X
	6 Controls (Histone H3, S6, GAPDH, Mouse IgG1, Mouse IgG2a, Rabbit IgG)	X	X	X
IO Drug Target Panel				
Human Protein Module	4-1BB	X	X	X
	ARG1	X	X	X
	B7-H3	X	X	X
	GITR	X	X	X
	IDO1	X	X	X
	LAG3	X	X	X
	OX40L	X	X	X
	STING	X	X	X
	TIM-3	X	X	X
	VISTA	X	X	X
Immune Activation Status Panel				
Human Protein Module	CD127	X	X	X
	CD25	X	X	X
	CD27	X	X	X
	CD40	X	X	X
	CD44	X	X	X
	CD80	X	X	X
	ICOS	X	X	X
	PD-L2	X	X	X

Cell Death Panel				
Human Protein Module	BAD	X		
	BCL6	X		
	BCLXL	X		
	BIM	X		
	CD95/Fas	X		
	GZMA	X		
	P53	X		
	PARP	X		
	Cleaved Caspase 9	X		
Immune Cell typing Panel				
Human Protein Module	CD14	X		
	CD163	X		
	CD34	X		
	CD45RO	X		
	CD66b	X		
	FAPalpha	X		
	FOXP3	X		
Pan-Tumor Panel				
Human Protein Module	BCL-2	X		
	EpCAM	X		
	ER alpha	X		
	HER2/ERBB2	X		
	MART1	X		
	NY-ESO-1	X		
	PR	X		
	PTEN	X		
	S100B	X		
MAPK Signaling Panel				
Human Protein Module	EGFR	X		
	Pan-RAS	X		
	BRAF	X		
	Phospho-c-RAF (S338)	X		
	Phospho-JNK (T183/Y185)	X		
	Phospho-MEK1 (S217/S221)	X		
	Phospho-p38 MAPK (T180/Y182)	X		
	Phospho-p44/42 MAPK ERK1/2 (T202/Y204)	X		
	P44/42 MAPK ERK1/2	X		
	Phospho-p90 RSK (T359/S363)	X		

Supplementary Table S3. Anti-CD44 antibody clone information

Antibodies	Source	Catalog #	Clone	Isotype	Concentration
CD44	Cell Signaling Technology (CST)	3570	156-3C11	mIgG2a	0.215 ug/ml
CD44	Abcam	ab51037	EPR1013Y	rIgG	0.192 ug/ml
Cytokeratin	DAKO	M351501-2	AE1/AE3	mIgG1	0.157 ug/ml
Cytokeratin	DAKO	Z062201-2	Polyclonal	rIgG	123 ug/ml

Supplementary Table S4. Association of CD44 expression measured by DSP in tumor and immune compartment with Overall Survival (OS) and Progression free Survival (PFS) in YTMA471 cohort

Overall survival (OS)						
Compartment	N	Marker	Univariate HR (95% CI)	P-value	Multivariate HR (95% CI)	P-value
Tumor compartment (panCK+)	54	CD44	0.82 (0.64-1.05)	0,13	NA	NA
Immune compartment (panCK-/CD45+/CD68+)	54	CD44	0.93 (0.58-1.48)	0,8	NA	NA
Progression-free survival (PFS)						
Compartment	N	Marker	Univariate HR (95% CI)	P-value	Multivariate HR (95% CI)	P-value
Tumor compartment (panCK+)	54	CD44	0.76 (0.61-0.96)	0,024	0.68 (0.46, 0.99)	0,043
Immune compartment (panCK-/CD45+/CD68+)	54	CD44	0.84 (0.55-1.28)	0,4	NA	NA

Supplementary tables S5. Impact of tumor cell CD44 expression in progression-free survival (PFS) in the univariate and bi-variate Cox Regression analysis after correcting for PD-L1 tumor proportion score (TPS) in the discovery set (a) and the validation set (b)

a)

Univariate Cox Regression analysis for PFS (YTMA471 cohort)			
Marker	**HR**	**CI 95%**	**p value**
PD-L1 IHC (TPS ≥1% vs. TPS<1%)	0.53	0.24-1.18	0.14
PD-L1 IHC (TPS ≥50% vs. TPS<50%)	0.48	0.21-1.09	0.064
Tumor CD44 QIF scores (highest quartile vs. rest)	0.29	0.11-0.75	0.004
Multivariate Cox Regression analysis for PFS (YTMA471 cohort)			
Co-variates	**HR**	**CI 95%**	**p value**
- PD-L1 IHC (TPS ≥1% vs. TPS<1%) - Tumor CD44 QIF scores (highest quartile vs. rest)	0.83 **0.25**	0.36-1.91 **0.09-0.69**	0.70 **0.003**
- PD-L1 IHC (TPS ≥50% vs. TPS<50%) - Tumor CD44 QIF scores (highest quartile vs. rest)	0.72 **0.26**	0.31-1.67 **0.10-0.72**	0.40 **0.004**

b)

Univariate Cox Regression analysis for PFS (H12O_ITX1 cohort)			
Marker	**HR**	**CI 95%**	**p value**
PD-L1 IHC (TPS ≥1% vs. TPS<1%)	0.91	0.62-1.34	0.66
PD-L1 IHC (TPS ≥50% vs. TPS<50%)	0.62	0.38-1.01	0.057
Tumor CD44 DSP counts (highest tertile vs. rest)	0.59	0.39-0.90	0.014
Multivariate Cox Regression analysis for PFS (H12O_ITX1 cohort)			
Co-variates	**HR**	**CI 95%**	**p value**
- PD-L1 IHC (TPS ≥1% vs. TPS<1%) - Tumor CD44 DSP counts (highest tertile vs. rest)	0.95 **0.59**	0.65-1.39 **0.39-0.90**	0.81 **0.015**
- PD-L1 IHC (TPS ≥50% vs. TPS<50%) - Tumor CD44 DSP counts (highest tertile vs. rest)	0.68 **0.62**	0.41-1.12 **0.41-0.96**	0.13 **0.032**

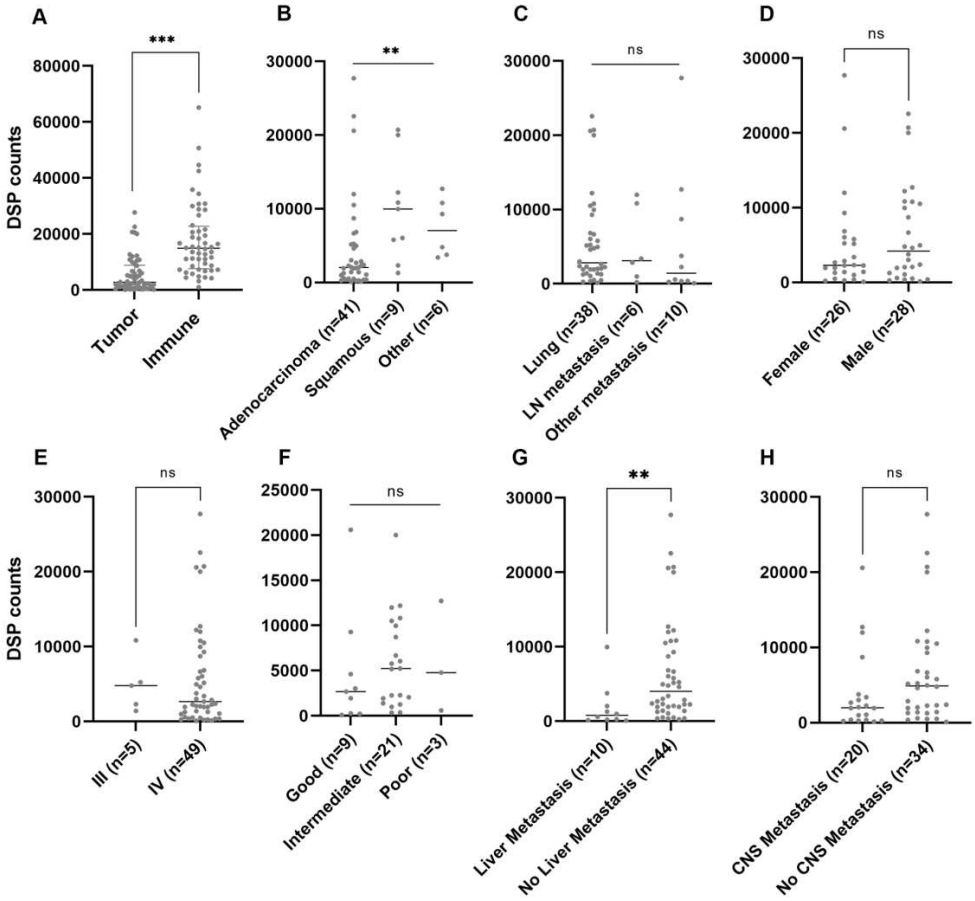
Supplementary Table S6. Correlation between CD44 expression and each individual protein marker all measured in the tumor compartment

H12O_ITX1			CIMA-CUN		
Protein	**R^2**	**P value**	**Protein**	**R^2**	**P value**
CD25	0,2432	<0,0001	CD80	0,3155	<0,0001
CD80	0,2038	<0,0001	CD40	0,3085	<0,0001
CD27	0,1429	<0,0001	PD-L1	0,2942	<0,0001
ICOS	0,1127	<0,0001	B7-H3	0,2934	<0,0001
GZMB	0,1071	<0,0001	ICOS	0,2316	<0,0001
CD40	0,09296	<0,0001	Tim-3	0,1968	<0,0001
PD-L2	0,0883	<0,0001	HLA-DR	0,1827	<0,0001
ARG1	0,06669	<0,0001	Beta-2-microglobulin	0,1406	<0,0001
CTLA4	0,058	<0,0001	CD27	0,1308	<0,0001
B7-H3	0,04519	<0,0001	CD3	0,08658	<0,0001
CD127	0,04186	<0,0001	CD8	0,08351	<0,0001
SMA	0,03677	<0,0001	GITR	0,08256	<0,0001
Tim-3	0,02972	<0,0001	VISTA	0,07979	<0,0001
Ki-67	0,02951	<0,0001	CD45	0,07487	<0,0001
LAG3	0,02208	0,0002	LAG3	0,06371	<0,0001
CD4	0,01726	0,0011	4-1BB	0,06155	<0,0001
VISTA	0,01655	0,0014	CD11c	0,05514	0,0002
PanCk	0,01638	0,0015	CD4	0,04975	0,0003
GITR	0,01404	0,0033	CD25	0,04683	0,0005
Fibronectin	0,01272	0,0052	STING	0,0458	0,0006
CD20	0,009499	0,0159	IDO1	0,04073	0,0012
4-1BB	0,007197	0,0359	GZMB	0,03753	0,0019
CD11c	0,006873	0,0403	PD-L2	0,0234	0,0145
CD68	0,005247	0,0733	CD68	0,02306	0,0152
CD45	0,004803	0,0867	PD-1	0,02232	0,017
CD8	0,003201	0,1621	CD127	0,01721	0,0363
PD-L1	0,002959	0,179	Fibronectin	0,01162	0,0858
CD3	0,001606	0,3223	PanCk	0,0107	0,0993
PD-1	0,001513	0,3367	Ki-67	0,008848	0,1341
STING	0,000521	0,5731	CD20	0,006468	0,2005
HLA-DR	0,000481	0,588	ARG1	0,003447	0,3504
CD56	0,000328	0,6549	SMA	0,003372	0,3558
Beta-2-microglobulin	0,000274	0,683	CTLA4	0,00151	0,5368
IDO1	0,000159	0,7553	OX40L	0,000829	0,6473
OX40L	2,08E-06	0,9716	CD56	0,000615	0,6934

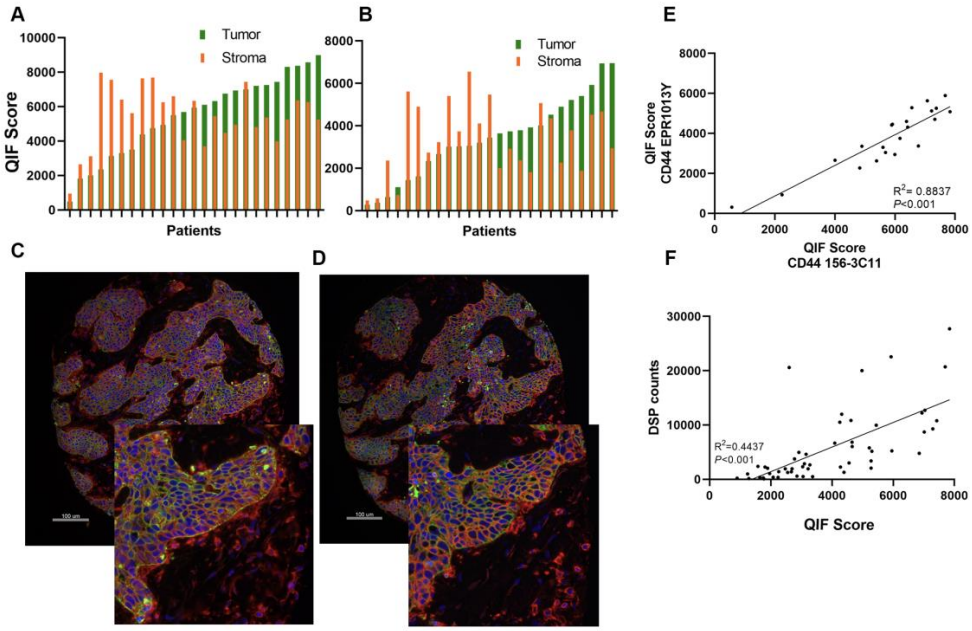
Supplementary Table S7. Differentially expressed protein markers between ROIs with elevated CD44 expression in the tumor compartment relative to ROIs with low CD44 expression in the tumor compartment

H12O_ITX1			CIMA-CUN		
Targets	**Log2foldChange (high-low)**	**FDR-adjusted p-value**	**Targets**	**Log2foldChange (high-low)**	**FDR-adjusted p-value**
PD-L1	1,2257	1,49E-08	Tim-3	0,80017	0,001035
IDO1	0,90745	0,0054572	CD40	0,79803	1,18E-05
VISTA	0,90323	8,54E-13	B7-H3	0,74146	1,18E-05
ICOS	0,88904	2,64E-13	ICOS	0,72723	1,24E-06
Tim-3	0,80696	1,09E-09	PD-L1	0,71736	0,011757
CD27	0,73234	1,49E-08	CD45	0,59472	0,002038
CD40	0,66583	2,11E-11	CD11c	0,57923	0,001035
ARG1	0,64844	2,11E-11	CD3	0,52081	0,003133
CD25	0,64208	7,42E-08	CD27	0,42725	0,002441
4-1BB	0,58834	9,08E-06	CD8	0,42541	0,006465
CD8	0,55807	1,61E-08	VISTA	0,38779	0,012837
CD45	0,53941	3,18E-06	Ki-67	0,21365	0,013243
B7-H3	0,52212	1,09E-09	IDO1	-0,42921	0,019737
GITR	0,4729	2,93E-09			
Ki-67	0,47068	9,32E-08			
Fibronectin	0,44567	0,00014619			
CD11c	0,44359	7,42E-06			
GZMB	0,37976	1,03E-10			
CD3	0,30722	0,0055482			
CD4	0,3037	1,06E-06			
CD68	0,22418	0,0043512			
PanCk	0,1781	0,00091907			
CD56	-0,32645	0,034899			

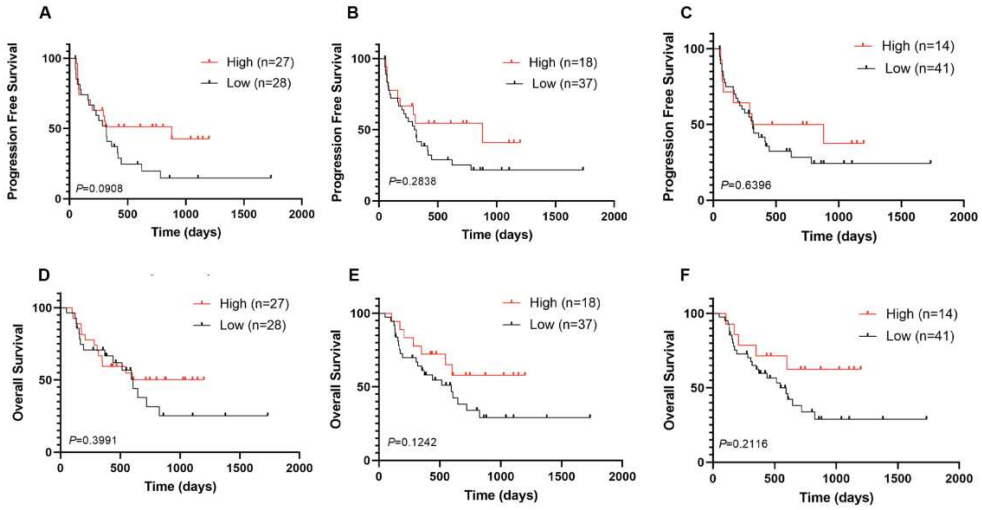
Only markers with a fold change (FC)≥1 and significant FDR-adjusted p-value (<0.05) are shown



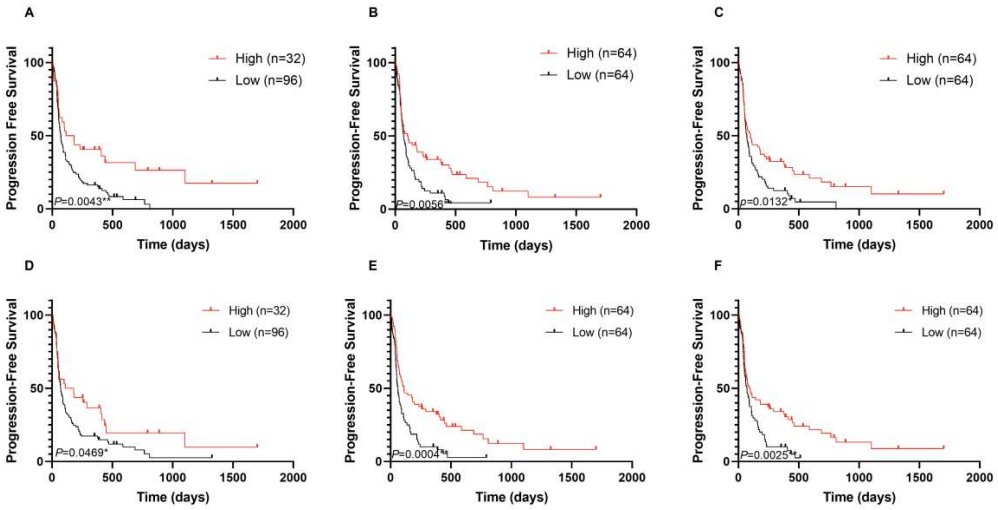
Supplementary Figure S1. **(A)** Comparative analysis of CD44 levels measured by DSP in the tumor compartment (panCK+) and in the immune compartment (panCK-/CD45+/CD68+) in YTMA471 cohort. **(B-H)** CD44 levels measured by DSP according to patients' clinical characteristics in YTMA471 cohort.
p-values legends *: P<0.05; **: P<0.01, ***:P<0.001, ns: not significant; LN: lymph node; CNS: central nervous system



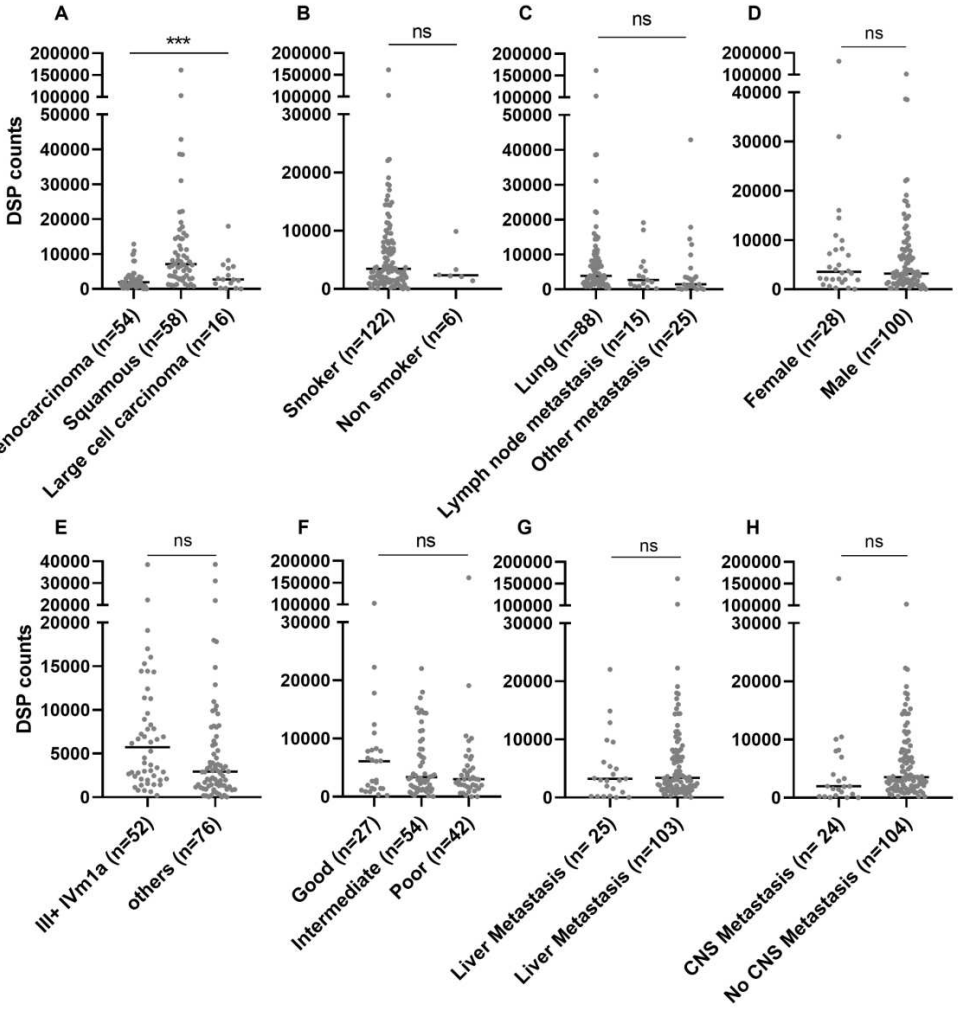
Supplementary Figure S2. (A-B). Dynamic range of CD44 QIF levels using two different anti-CD44 clones in a control lung cancer array (YTMA295); CD44 156-3C11 (A) and CD44 EPR1013Y (B). **(C-D).** Representative images of CD44 expression in the same TMA spot using CD44 156-3C11 (C) and CD44 EPR1013Y (D). **(E)** Correlation analysis of QIF scores obtained with the two different CD44 clones in YTMA295. **(F)** Correlation analysis of CD44 quantitative measurements obtained by DSP counts and QIF scores in YTMA471.



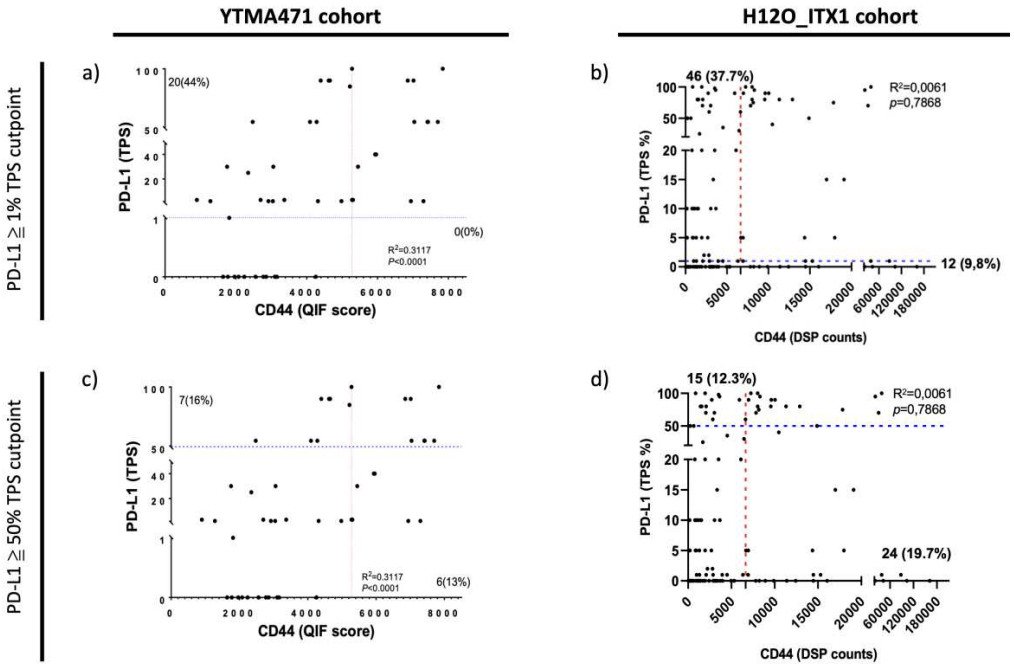
Supplementary Figure S3. **(A-C)** Kaplan-Meier PFS curves according to CD44 expression (QIF scores) in the immune compartment using the median (A), tertile (B), and quartile (C) cutpoints in YTMA471 cohort. **(D-F)** Kaplan-Meier OS curves according to CD44 expression (QIF scores) in the immune compartment using the median (D), tertile (E), and quartile (F) cutpoints in YTMA471 cohort.



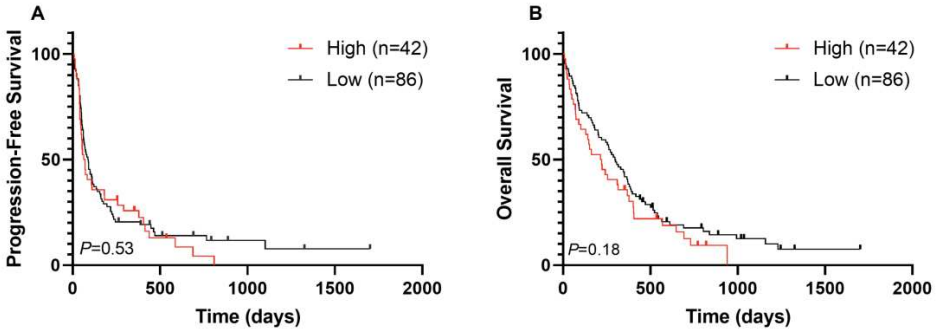
Supplementary Figure S4. (A-C) Kaplan-Meier PFS curves according to PD-L1 (A), CD3 (B), and CD8 (C) expression in the tumor compartment in H120_ITX1 cohort. **(D-F)** Kaplan-Meier PFS curves according to PD-L1 (D), CD3 (E), and CD8 (F) expression in the immune compartment in H120_ITX1 cohort.



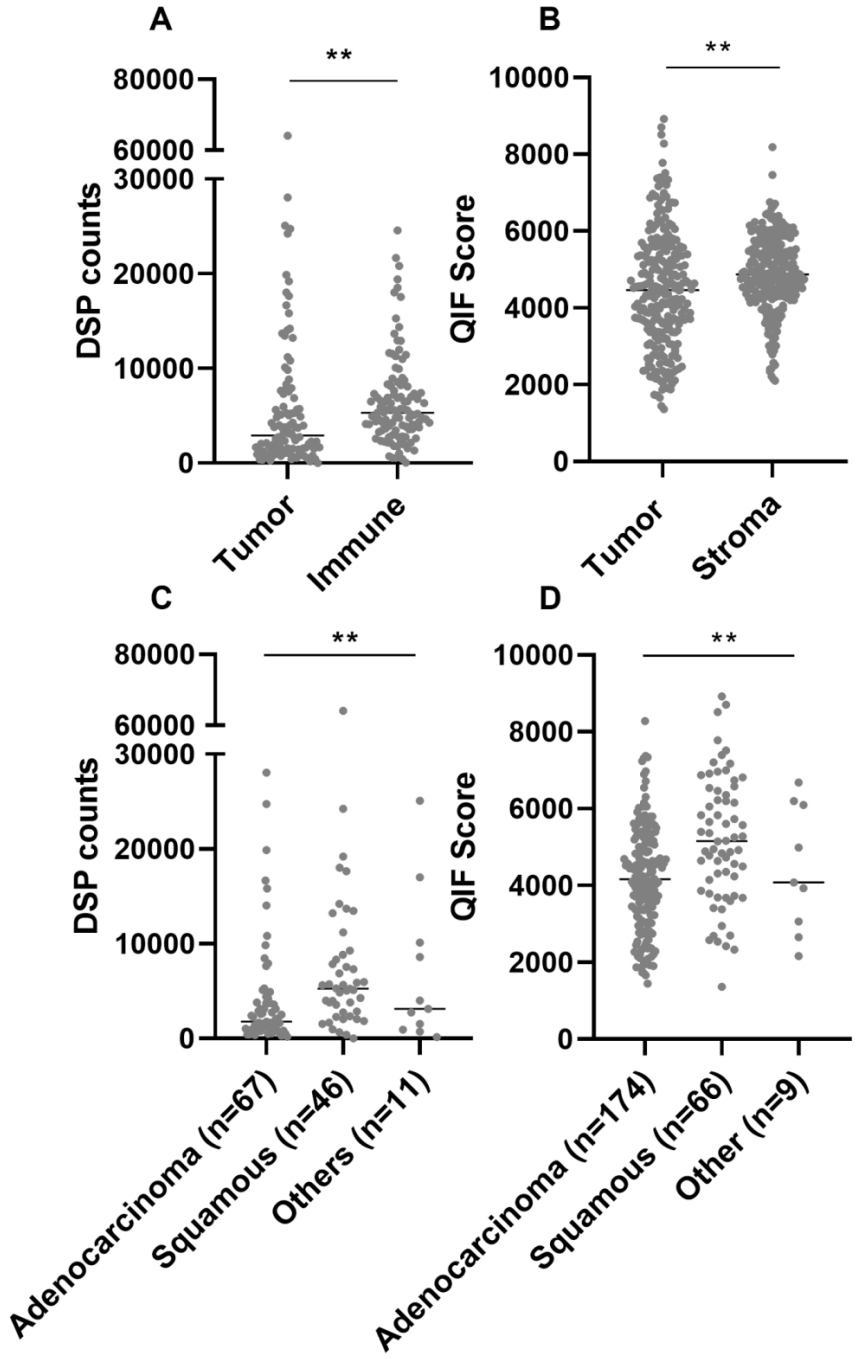
Supplementary Figure S5. CD44 levels measured by DSP according to patients' clinical characteristics in H12O_ITX1 cohort.
p-values legends *: P<0.05; **: P<0.01, ***:P<0.001, ns: not significant



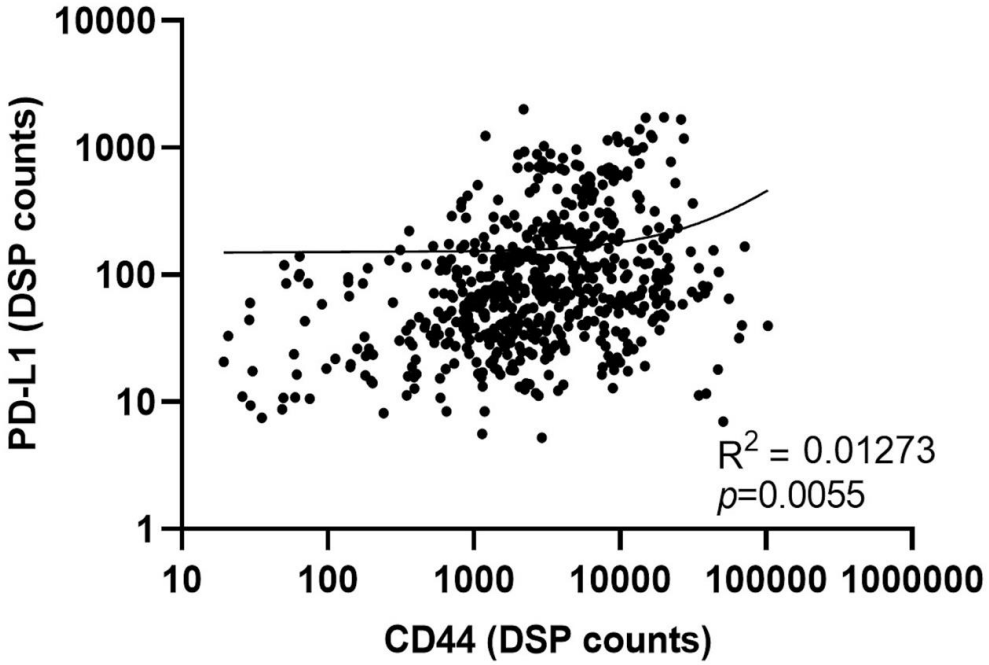
Supplementary figure S6. Correlation charts illustrating the association of tumor cell CD44 and PD-L1TPS in the YTMA471 and H12O_ITX1 cohorts. Red line highlights tumor cell CD44 cutpoints, blue line highlights PD-L1 TPS cutpoints. **(A-B)** Graphs illustrating the number (%) of discordant cases using the PD-L1≥1% TPS cutpoint in both cohorts. **(C-D)** Graphs illustrating the number (%) of discordant cases in both cohorts using the PD-L1≥50% TPS cutpoint in both cohorts.



Supplementary Figure S7. (A-B) Kaplan-Meier PFS (A) and OS (B) curves according to CD44 expression in the immune compartment using the top tertile cutpoint in H12O_ITX1 cohort.



Supplementary Figure S8. (A-B) Comparative analysis of CD44 levels in the tumor and immune compartments using DSP in CIMA-CUN cohort (A) and tumor and stroma compartments using QIF in YTMA423 cohort (B). **(C-D)** CD44 levels according to histology measured by DSP in CIMA-CUN cohort (C) and by QIF in YTMA423 cohort (D).
p-values legends *: P<0.05; **: P<0.01, ***:P<0.001, ns: not significant



Supplementary figure S9. Correlation chart showing the association between tumor cell CD44 and PD-L1 in the H12O_ITX1 cohort after removing outlier values exceeding the upper limits of expression of tumor cell CD44 and PD-L1 in the CIMA-CUN cohort