Int J Clin Exp Pathol 2026;19(6):235-247
www.ijcep.com /ISSN:1936-2625/IJCEP0172156

Original Article

A three-protein serum risk score for predicting immunotherapy response and prognosis in non-small cell lung cancer

Yiming Ma[1*], Yuan Gao[2*], Changjian Shao[1*], Kaiqi Wei[1], Wenchen Wang[1], Qiongjie Shao[1], Tao Jiang[1]

[1]*Department of Thoracic Surgery, Tangdu Hospital, Air Force Medical University, Xi'an 710038, Shaanxi, China;* [2]*The State Key Laboratory of Cancer Biology, Biotechnology Center, School of Pharmacy, Air Force Medical University, Xi'an 710032, Shaanxi, China.* [*]*Equal contributors.*

Received February 16, 2026; Accepted April 27, 2026; Epub June 15, 2026; Published June 30, 2026

Abstract: Background/Objective: Immune checkpoint inhibitors (ICIs) have extended survival in patients with non-small cell lung cancer (NSCLC) but their therapeutic benefit is limited to a proportion of patients. Predictive biomarkers based on tissue of origin of the tumor have their limitations, and thus there is a need for solid and minimally invasive predictive biomarkers. Our aim was to investigate serum proteomics via liquid biopsy for biomarker discovery. Methods: In this retrospective extension of the TD-FOREKNOW trial, deep proteomic profiling was undertaken on pre-treatment serum samples of 72 patients with NSCLC receiving neoadjuvant therapy. Further quantitation of proteins in serum was performed by data independent acquisition mass spectrometry to obtain candidates associated with treatment outcome. Statistical regression was also used to screen for proteins related to ICI efficacy and a risk score composite model was set up to predict treatment response and prognosis. Results: From the 1,802 analyzed serum proteins, 59 serum proteins were differentially expressed in patients receiving immunotherapy plus chemotherapy. Using univariate logistic regression followed by least absolute shrinkage and selection operator (LASSO) regression, three factors, SERPINE2, DAZAP1, and MGAT4B, were identified whose baseline expression was correlated with the response to ICI therapy. The risk score model using the three proteins was an effective biomarker in predicting ICI response with an area under the curve (AUC) of 0.946 (95% CI: 0.874-1.000). Its predictive value for ICI response was validated in further survival analysis, showing that patients with a low risk score had significantly longer progression-free survival (HR = 0.13, 95% CI: 0.04-0.46, P = 0.002) and overall survival (HR = 0.14, 95% CI: 0.03-0.62, P = 0.033) than those with a high risk score. Conclusion: A pre-treatment serum-based risk score that effectively predicts response to ICI therapy in patients with NSCLC was developed and validated. Our findings reveal the great prospect of human serum proteomics as a powerful liquid biopsy platform for biomarker discovery and construction of clinical prognostic models.

Keywords: Immunotherapy, efficacy, non-small cell lung cancer, prediction model, proteomics

Introduction

In China, lung cancer is the second most common cancer and the leading cause of cancer death [1]. Five-year survival rate after diagnosis is only 10-20% [2]. The arrival of immune checkpoint inhibitors (ICIs) changed this landscape. Monoclonal antibodies targeting programmed cell death protein 1 (PD-1) or programmed cell death ligand 1 (PD-L1) have fundamentally altered how we treat locally advanced non-small cell lung cancer (NSCLC) [3]. Their mechanism of action lies in relieving tumor-induced suppression of T cells and reactivating the body's antitumor immune response. Compared to chemotherapy alone, chemoimmunotherapy significantly improves the pathological complete response (pCR) rate in patients with resectable NSCLC. Studies such as CheckMate 816 and KEYNOTE-671 have confirmed this potential [4, 5]. Our earlier TD-FOREKNOW trial also reached similar conclusions [6]. Meta-analyses of randomized trials have demonstrated superior event-free survival and overall survival for the chemoimmunotherapy arm, confirming its long-term ben-

https://doi.org/10.62347/EMOY9026

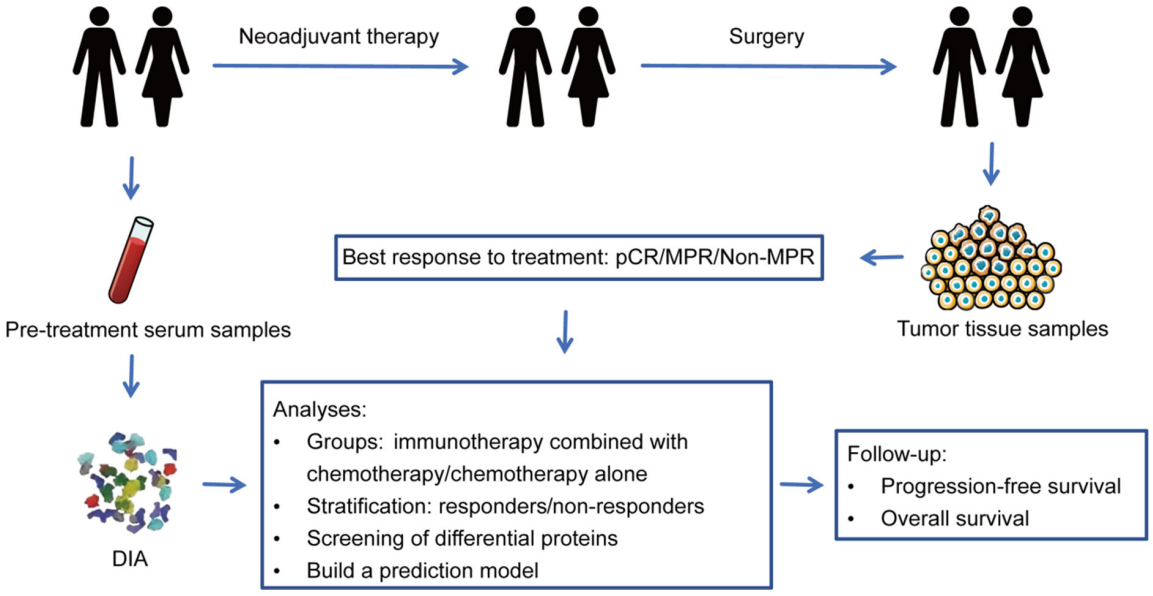

Figure 1. Study workflow. DIA, data-independent acquisition; pCR, pathological complete response; MPR, major pathological response; Non-MPR, non-major pathological response.

efit over chemotherapy alone [3]. In summary, chemoimmunotherapy demonstrates unique advantages in both short-term efficacy evaluation and long-term survival benefits. However, most lung cancer patients remain unlikely to achieve a meaningful response to ICIs. The response rate is less than 50% in the first-line setting and merely 20% in the second-line setting [7, 8].

Current biomarkers for optimization of patient selection include tumor mutation burden (TMB) [9], PD-L1 expression [10], and ENLIGHT-DP [11]. Most of these biomarkers show suboptimal predictive efficacy and their use is constrained by poor accessibility and temporal-spatial heterogeneity. This highlights the importance of identifying novel biomarkers to accurately select patients who will benefit from this impactful but costly therapy.

Peripheral blood biomarkers offer a non-invasive means to profile the tumor immune microenvironment. These include exosomes [12], circulating tumor DNA [13], cell-free DNA [14], cell-free RNA [15], and peripheral blood proteins. Among these, peripheral blood proteins play an "executive" role in cellular physiological functions and may reflect the tumor immune microenvironment. Plasma-based proteins have also been used as biomarkers to predict prognosis and monitor treatment efficacy. A recent large-scale data-independent acquisition (DIA) proteomics study in prostate cancer identified a 16-protein plasma panel that effectively predicts biochemical recurrence [16]. Compared with traditional biomarkers, such as PD-L1 expression status, which are constrained by temporal and spatial heterogeneity, circulating biomarkers represent a promising alternative to address the limitations associated with tumor tissue analysis. Thus, plasma proteomics represents a powerful and novel platform for discovering clinically actionable, protein-based biomarkers. This study was carried out with the objective of exploring serum proteomics for biomarker discovery to optimize patient selection for immunotherapy.

Materials and methods

The study design is shown in **Figure 1** and is briefly described below.

Study population and data collection

This retrospective study is an analytical extension of the TD-FOREKNOW trial (ClinicalTrials.gov Identifier: NCT04338620). The TD-FOREKNOW trial was a randomized, multicenter, phase 2 clinical trial and is already published. In this trial, eligible participants with resectable stage IIIA or IIIB (N2) NSCLC were random-

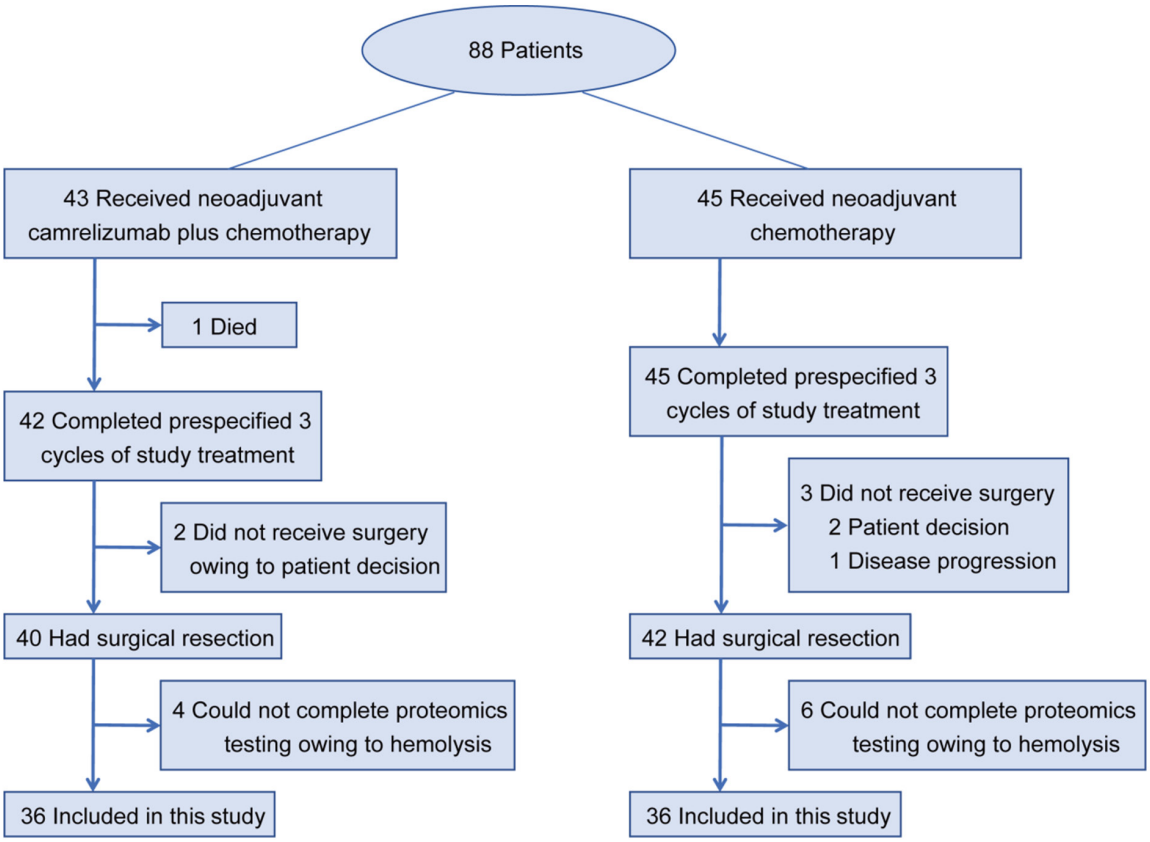

Figure 2. Selection of patients based on their clinical history.

ized (1:1) to receive either three cycles of neoadjuvant camrelizumab plus nab-paclitaxel and platinum-based agents or three cycles of chemotherapy alone, to evaluate the efficacy and safety of neoadjuvant camrelizumab combined with chemotherapy compared with neoadjuvant chemotherapy alone [6].

Based on the patient cohort studied in the TD-FOREKNOW trial, we enrolled patients according to the following inclusion and exclusion criteria to ensure the subsequent proteomics analysis could be conducted:

Inclusion criteria: 1) Patients belonging to the cohort of the TD-FOREKNOW trial; 2) Completion of all cycles of neoadjuvant therapy as per the trial protocol; 3) Completion of surgery with pathological evaluation results available for surgical specimens, whether pCR or major pathological response (MPR); 4) Well-preserved serum samples collected prior to neoadjuvant therapy that passed quality control before proteomics testing.

Exclusion criteria: 1) Failure to complete all cycles of neoadjuvant therapy; 2) Absence of surgical intervention or postoperative pathological evaluation results; 3) No serum samples collected prior to neoadjuvant therapy, or samples that failed quality control prior to proteomics testing.

According to the inclusion and exclusion criteria outlined above, a total of 72 patients were enrolled in this study, with 36 cases in each group (**Figure 2**).

This retrospective analysis of anonymous data was approved by the IEC of Institution for National Drug Clinical Trials of Tangdu Hospital (Approval No. K202406-18) with a waiver of informed consent.

Clinical data collection

Patients' clinical data included age at treatment initiation, sex, Eastern Cooperative Oncology Group (ECOG) stage, pathology type,

pathologic stage, best treatment response, date of disease progression, and/or date of death. These data were derived from our TD-FOREKNOW trial. The ECOG score was performed within 7 days prior to the first dose by the attending physician based on the patient's physical condition using the standard ECOG scoring scale. The pathological type was determined according to the pathology report performed within 90 days before the first dose. Pathological staging was performed according to the eighth edition of the clinical TNM staging of tumors. Pathological response was determined based on the results of pathologic testing of surgical specimens as pCR or MPR. pCR was defined as complete absence of viable tumor cells in the primary tumor specimen and all sampled regional lymph nodes. MPR was defined as the presence of ≤10% live tumor cells in the resected primary tumor specimen and sampled regional lymph nodes. Responders were defined as having MPR, and non-responders as failure to achieve MPR. The date of progression was recorded at follow-up. Progression-free survival (PFS) was defined as time from treatment initiation to disease progression or death. Overall survival (OS) was defined as the time from treatment initiation to death.

Blood sample collection and storage

All blood samples were collected in serum-separating plain tubes prior to neoadjuvant therapy. Following collection, they were incubated at 37°C for 2 h and then centrifuged at 3,000×*g* for 10 min to separate plasma. The separated plasma was further centrifuged at 1,800×*g* for 8 min and stored at -80°C until further analysis.

Serum protein expression assay

Protein expression assay was performed at the Clinical Biomarker Facility of Shanghai Aki Biotechnology Co., Ltd. Proteolytic peptides were analyzed via liquid chromatography-tandem mass spectrometry (LC-MS/MS, timsTOF Pro2, Bruker, Germany) using DIA mode and Spectronaut software (version 18.2.230802.50606; Biognosys AG) to identify and quantify the proteome and to obtain protein expression information. All the samples showing proteins with ≥1 unique peptide were retained. Missing-value recoding (MVR) on the original data was performed using DIA mode and imputed using the half-minimum method.

Sample measurements were 25 μL each, which was protein-enriched, proteolyzed, and peptide desalted, and subsequently detected via nanoLCMS/MS (timsTOF Pro2, Bruker, Germany). Mass spectrometry analysis was performed in data independence acquisition with parallel accumulation-serial fragmentation (DIA-PASEF) mode with a scanning range of 100-1,700 m/z. During PASEF MS/MS scanning, collision energies increased linearly with ion mobility from 20 eV (1/K0 = 0.85 Vs/cm^2) to 59 eV (1/K0 = 1.30 Vs/cm^2). Raw mass spectrum files were imported into Spectronaut (version 18.2.230802.50606; Biognosys AG) for database searching and qualitative analysis.

Differential protein screening

Differentially expressed proteins were screened in two groups, viz. chemotherapy and immunotherapy-combined with chemotherapy (immunochemotherapy) groups based on response status. The preliminary screening criteria for efficacy-based differential proteins within the two groups were: 1) *P*-value <0.05 based on Student's t-test or chi-square test; 2) FOLD CHANGE ≤0.83 or FOLD CHANGE ≥1.2. In the chemotherapy group, a total of 337 differentially expressed proteins were identified, among which 47 were upregulated and 290 were downregulated in responders compared to non-responders. In the immunochemotherapy group, 60 differentially expressed proteins were identified, with 41 upregulated and 19 downregulated. To identify the potential predictive protein biomarkers for immunotherapy efficacy, 59 proteins showing significantly different expression in the immunochemotherapy group were selected.

Construction of a risk score prediction model for immunotherapy efficacy

Raw data for protein expression were first normalized by removing proteins with >30% missing values. The remaining missing values were imputed using 1/2 of the smallest expression value for each protein. Thereafter, univariate logistic regression was performed for each pro-

tein showing a *p*-value of <0.05 using R software. Subsequently, a least absolute shrinkage and selection operator (LASSO) regression was applied to prevent overfitting and select the most predictive features. This yielded a final model based on three proteins: SERPINE2, DAZAP1, and MGAT4B. The resulting risk score was defined as: Risk Score = -1.365144071 + (0.003770902) × SERPINE2 + (0.027029983) × DAZAP1 + (-0.021299981) × MGAT4B. The discriminatory performance of this model was evaluated by generating a receiver operating characteristic (ROC) curve and area under the ROC curve (AUC) using GraphPad Prism. The generalization ability of the model was evaluated using the hierarchical bootstrap method.

Prognostic analysis

To evaluate the prognostic value of the identified biomarkers, we calculated a risk score for each patient using the previously established three-protein model. Patients were then stratified into high-risk and low-risk groups based on the median risk score. Survival outcomes, PFS and OS were compared between these groups using Kaplan-Meier curves with log-rank tests. Univariate Cox proportional hazards regression was applied to the full cohort to further assess the association between the risk score and survival outcomes.

Statistical analyses

Categorical variables are expressed as proportions and comparisons between two categorical variables were performed using Fisher's exact test or chi-square test (as appropriate). For the chemotherapy group or the immunotherapy combined with chemotherapy group, differential proteins were preliminarily screened based on therapeutic efficacy using the Student's t-test (P<0.05) or chi-square test (P<0.05), with FOLD change criteria of ≤0.83 or ≥1.2. Model construction employed univariate logistic regression (P<0.05) and LASSO regression. Model validation utilized ROC curves and the stratified bootstrap method. Survival analysis was conducted using Kaplan-Meier curves combined with the log-rank test, with validation performed through univariate Cox regression analysis. Statistical analyses were performed using GraphPad Prism (V.10.0), SPSS (V.23), and R (V.4.4.3) statistical software. Statistical significance was defined as a two-tailed P<0.05.

Results

Characteristics of study population

Among the total 72 patients with NSCLC enrolled in this study, 36 received neoadjuvant chemotherapy alone and 36 received neoadjuvant immunochemotherapy. Among them 47 patients (65.3%) were diagnosed with lung squamous cell carcinoma whereas, 22 (30.6%) were diagnosed with lung adenocarcinoma. According to the definition set for this study, 30 patients were identified as responders, among which five belong to the chemotherapy group and 25 to the immunochemotherapy group. Forty-two patients were identified as non-responders, with 31 in the chemotherapy group and 11 in the immunochemotherapy group. Demographic and clinical characteristics of the enrolled patients did not show significant differences between responders and non-responders (**Table 1**).

Identification of serum proteins associated with immunotherapy efficacy

To identify serum protein biomarkers predictive of response to ICI therapy, we performed differential expression analysis across the 1,802 quantified proteins. Comparative analysis between responders and non-responders revealed 337 (47 upregulated and 290 downregulated) differentially expressed proteins (DEPs) in the chemotherapy-alone cohort, and 60 DEPs (41 upregulated and 19 downregulated) in the immunochemotherapy cohort (P<0.05; FOLD change ≤0.83 or ≥1.2) (**Figure 3**).

To specifically isolate a proteomic signature of ICI efficacy, we filtered for proteins uniquely associated with immunochemotherapy response. This yielded a refined set of 59 candidate biomarkers. These proteins were either exclusively differentially expressed in the immunochemotherapy cohort or exhibited an opposite expression trend (i.e., upregulated vs. downregulated) compared to the chemotherapy-alone cohort (**Figure 4**).

Table 1. Baseline clinical characteristics of patients

Patient characteristics	Overall (n = 72)	Non-Responder (n = 42)	Responder (n = 30)	*P* value
Age (%)				
<60	29 (40.3)	17 (40.3)	12 (40.3)	0.968
≥60	43 (59.7)	25 (59.7)	18 (59.7)	
Gender (%)				
Male	62 (86.1)	38 (90.5)	24 (80.0)	0.357
Female	10 (13.9)	4 (9.5)	6 (20.0)	
ECOG (%)				
0	71 (98.6)	41 (97.6)	30 (100.0)	1.000
1	1 (1.4)	1 (2.4)	0 (0.0)	
Treatment (%)				
Immunotherapy and chemotherapy	36 (50.0)	11 (26.2)	25 (83.3)	<0.001
Chemotherapy	36 (50.0)	31 (73.8)	5 (16.7)	
Pathological type (%)				
Squamous	47 (65.3)	25 (59.5)	22 (73.3)	0.523
Adenocarcinoma	22 (30.6)	15 (35.7)	7 (23.3)	
Other	3 (4.2)	2 (4.8)	1 (3.3)	
Stage (%)				
IIIA	52 (72.2)	31 (73.8)	21 (70.0)	0.722
IIIB	20 (27.8)	11 (26.2)	9 (30.0)	

Note: Values outside parentheses in the table represent variable counts, while those within parentheses indicate proportions. Comparisons between two categorical variables were performed using Fisher's exact test or chi-square test (as appropriate).
Abbreviation: ECOG, Eastern Cooperative Oncology Group.

Development and validation of a risk score for immunotherapy response

To develop a robust predictor of ICI efficacy, the 59 candidate proteins were subjected to further variable selection. Using univariate logistic regression, we found 18 proteins whose expression levels correlated with treatment outcome (P<0.05). LASSO regression then narrowed these down to three key proteins: SERPINE2, DAZAP1, and MGAT4B. We defined the final risk score model as: Risk Score = -1.365 + (0.003771 × SERPINE2) + (0.02703 × DAZAP1) + (-0.02130 × MGAT4B).

The composite risk score predicted immunotherapy response with high accuracy, achieving an AUC of 0.946 (95% CI: 0.874-1.000). By comparison, when we applied the same score to the chemotherapy-alone cohort, the AUC dropped to 0.613 (95% CI: 0.380-0.846). The three-protein model performed better than any single protein alone. For comparison, the AUCs for SERPINE2, DAZAP1, and MGAT4B used individually were 0.800, 0.867, and 0.778, respectively (**Figure 5**). Hierarchical bootstrap validation (1,000 iterations) confirmed the model's robustness, yielding a similar AUC of 0.946 (95% CI: 0.855-1.000) (**Figure 6**).

The risk score prediction model can predict progression-free and overall survival

To explore the prognostic significance of the three-protein risk score model, correlation analysis was performed in the full patient cohort. Patients were stratified into high-risk and low-risk groups based on the median risk score. Kaplan-Meier analysis showed that the risk score was significantly linked to both PFS and OS. Patients in the low-risk group demonstrated significantly longer PFS (HR = 0.13, 95% CI: 0.04-0.46, log-rank P = 0.002) and OS (HR = 0.14, 95% CI: 0.03-0.62, log-rank P = 0.033) than those in the high-risk group (**Figure 7**). Univariate Cox regression confirmed that a low-risk score was independent predictor of both longer PFS (HR = 0.16, 95% CI: 0.04-0.57, P = 0.005) and longer OS (HR = 0.17, 95% CI: 0.04-0.76, P = 0.021). We further explored the individual contributions of the model components to survival. Lower baseline expression of DAZAP1 was significantly associated with longer OS (log-rank P = 0.033), while high-

Figure 3. Comparative analysis of serum proteins differentially expressed by treatment response in each cohort. Venn diagrams categorize proteins that were significantly upregulated (high) or downregulated (low) in responders versus non-responders within the chemotherapy-alone and immunochemotherapy groups. A. Proteins downregulated in the immunochemotherapy responders but upregulated in the chemotherapy responders (0 overlapping proteins). B. Proteins upregulated in the immunochemotherapy responders but downregulated in the chemotherapy responders (7 overlapping proteins). C. Proteins downregulated in responders in both treatment groups (1 overlapping protein). D. Proteins upregulated in responders in both treatment groups (0 overlapping proteins). Overlapping proteins represent shared response signatures, while non-overlapping proteins indicate treatment-specific biomarkers. For the chemotherapy group or the immunotherapy combined with chemotherapy group, differential proteins were preliminarily screened based on therapeutic efficacy using Student t-test ($P<0.05$) or chi-square test ($P<0.05$), with FOLD change criteria of ≤ 0.83 or ≥ 1.2.

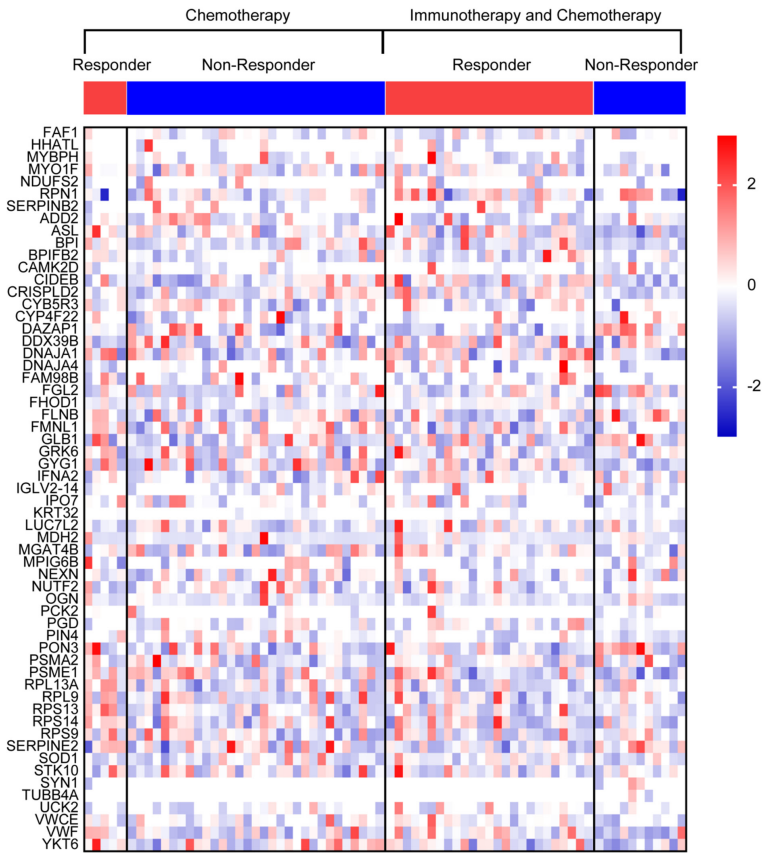

Figure 4. Distinct serum proteomic signatures differentiate treatment response in the immunochemotherapy cohort. The heatmap displays the expression levels of the 59 candidate biomarker proteins selected for their specific association with immunotherapy response. In the immunochemotherapy group, clear differences in protein expression are evident between responders and non-responders. In contrast, the same proteins either show no differential expression or exhibit an opposite expression trend in the chemotherapy-alone group. Each column represents an individual patient, and each row represents a single protein.

er baseline expression of MGAT4B was significantly associated with longer PFS (log-rank P = 0.037) (**Figure 8**).

Discussion

In this retrospective biomarker study, we were able to utilize high-depth serum proteomics to discover and validate a new three-protein signature-SERPINE2, DAZAP1 and MGAT4B-that accurately predicts response to neoadjuvant immunochemotherapy and stratifies survival in patients with NSCLC. To our knowledge, this was the first study to use DIA proteomics in serum sample to predict an immunotherapy outcome in patients with NSCLC, and the risk score model derived from it constructed a very accurate model (AUC, 0.946), highlighting its potential application as a clinical 'liquid biopsy' test for patients.

SERPINE2 (also known as the protease linker protein 1 PN-1) also emerged as a major predictive protein; higher baseline serum expression of SERPINE2 was associated with poor response and poor prognosis. Besides promoting tumor progression, SERPINE2 has been shown to correlate with poor outcomes in lung adenocarcinoma, hepatocellular carcinoma, and renal cell carcinoma [17-19]. Mechanistically, knocking down SERPINE2 inhibited radioresistance and DNA damage repair by regulating RAD51 in lung cancer. SERPINE2 interacted with ATM directly, which facilitated phosphorylation of ATM and homologous recombinational repair for DNA damage. In terms of immunotherapy, we speculate that SERPINE2 may increase the immunosuppressive tumor microenvironment by enhancing the survival of tumor cells and DNA repair capacity of tumor cells, thereby lowering immunogenicity and resistance to immunotherapy agents (ICI).

DAZAP1 is an RNA-binding protein with aberrant expression linked to proliferation, metastasis and prognosis in colorectal cancer, gastric cancer and melanoma [20-22]. Most recently, in melanoma that is unresponsive to anti-PD-1 treatment, researchers reported that a cancer-specific lncRNA named LISRR has high expression. Under treatment pressure, LISRR recruits DAZAP1 to ribosomes of the endoplasmic reticulum to promote synthesis of an immunosuppressive translatome that helps tumor cells escape from T-cell recognition and attack. The translatome comprised two major elements: PD-L1 and glycocalyx syn-

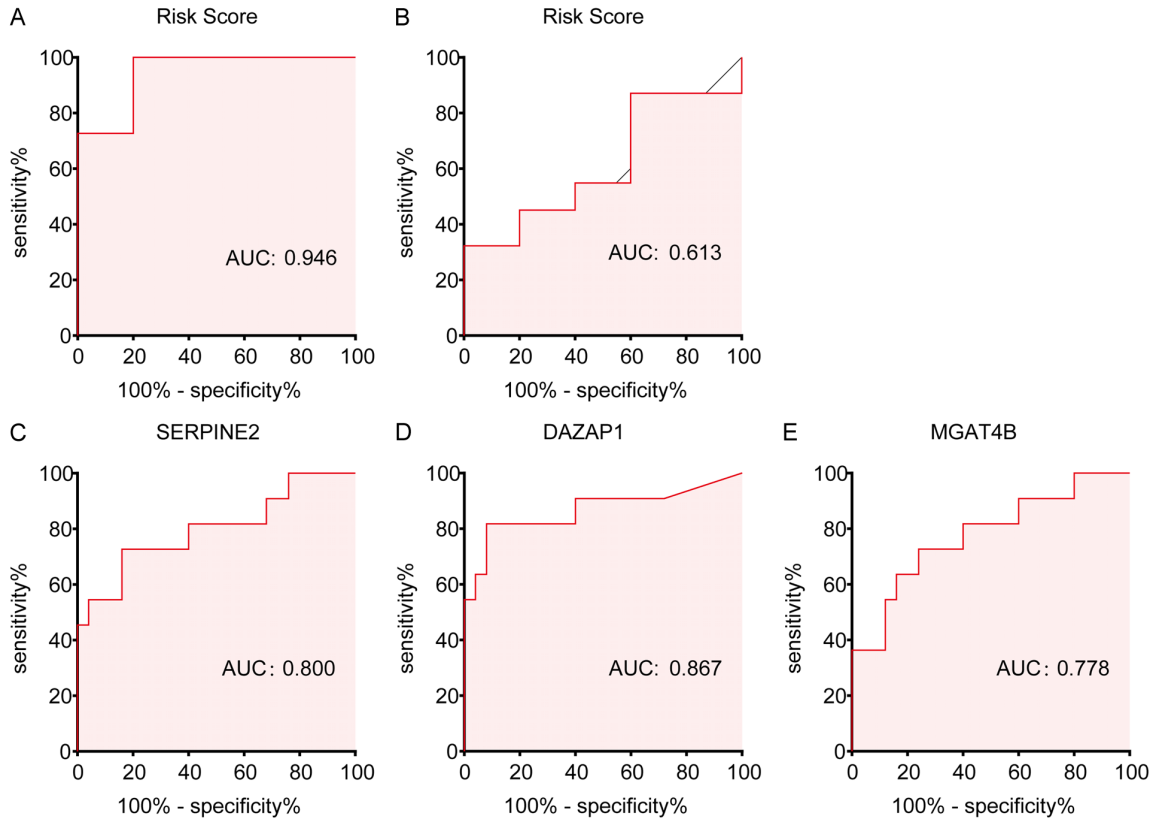

Figure 5. Receiver operating characteristic (ROC) curve analysis of the predictive model. (A) ROC analysis of the composite three-protein risk score for predicting treatment response in the immunochemotherapy cohort. (B) Application of the same risk score to the chemotherapy-alone cohort. (C-E) ROC curves for the individual model components-SERPINE2 (C), DAZAP1 (D), and MGAT4B (E) when used as single biomarkers within the immunochemotherapy cohort. The area under the curve (AUC) values demonstrate the superior predictive accuracy of the composite model compared to its individual constituents.

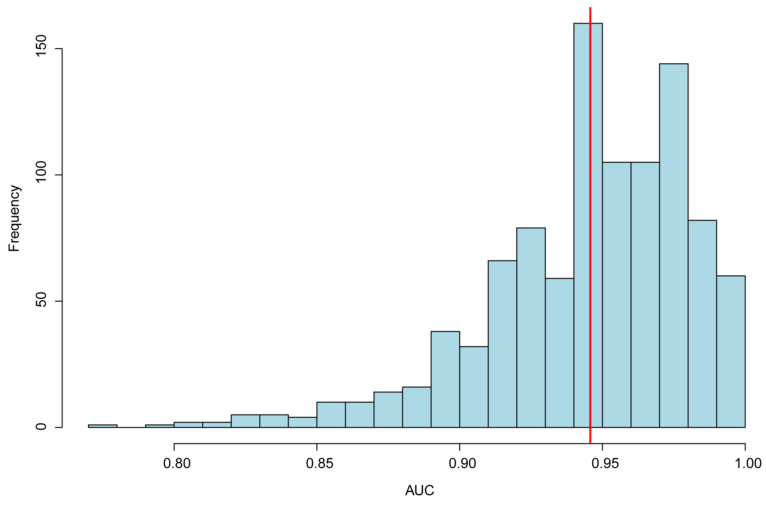

Figure 6. Hierarchical bootstrap AUC distribution.

thases. From these findings, we could reasonably assume that in patients with NSCLC undergoing neoadjuvant immunotherapy, under pressure of treatment, tumor cells may be activated by a similar LISRR-DAZAP1 axis to upregulate PD-L1 and remodel surface glycocalyx structure of tumor cells, so as to establish an immunosuppressive tumor environment. This could open up another possibility for exploring the role of RNA-binding proteins in regulating antitumor immunity.

MGAT4B is N-acetylglucosaminyltransferase-IVb and catalyzes β1-4 branching modifications of N-glycans on cell surface receptors, including integrins and growth factor receptors. These N-glycans carry

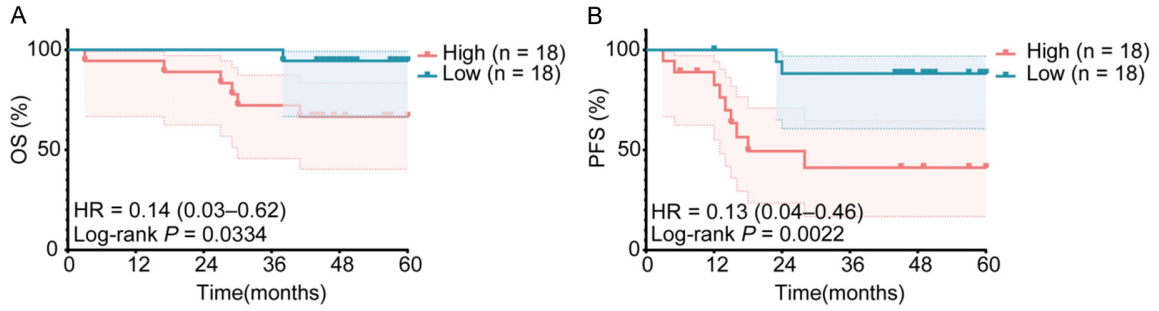

Figure 7. Kaplan-Meier survival curves of patient subgroups stratified by model risk score. OS, overall survival; PFS, progression-free survival.

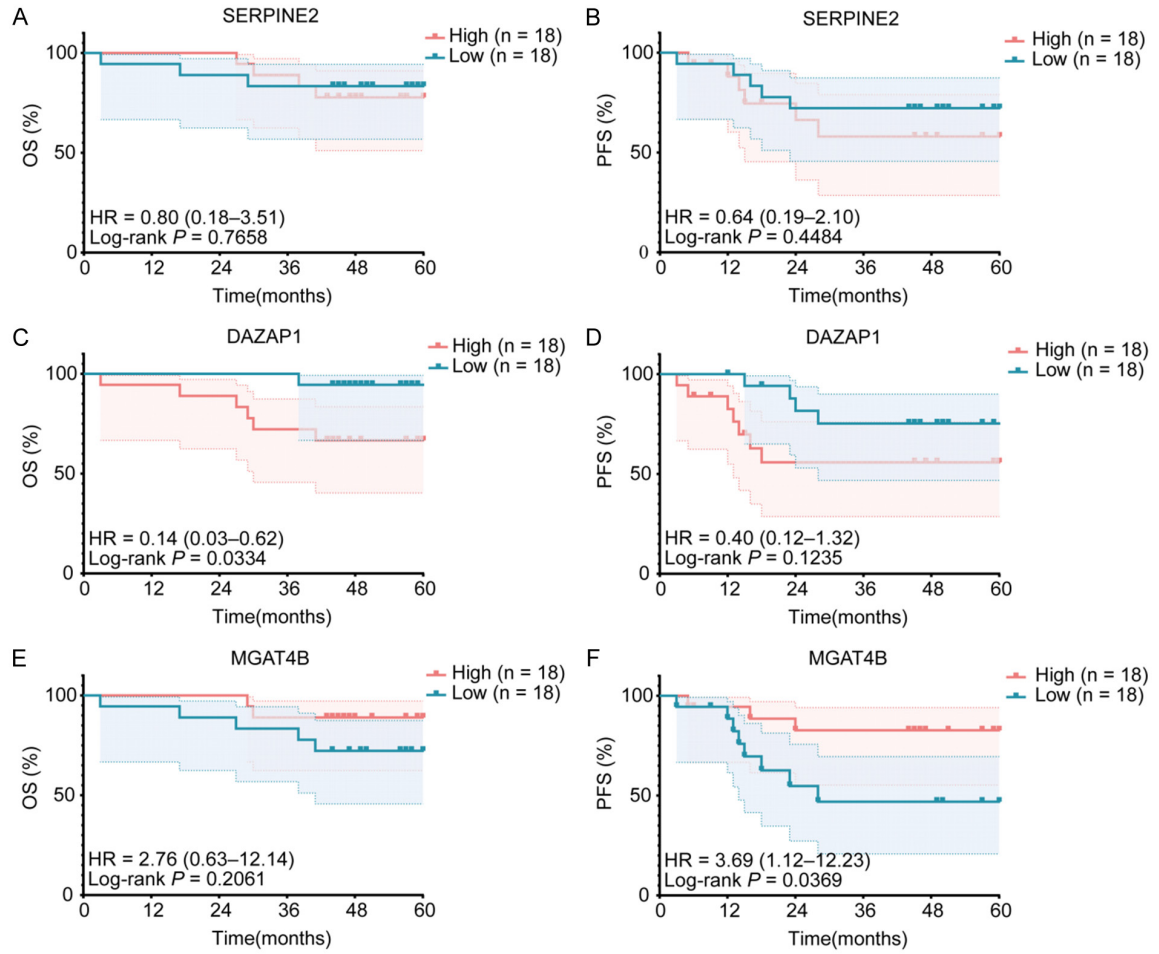

Figure 8. Kaplan-Meier survival curves for patients stratified by expression levels of the three proteins. OS, overall survival; PFS, progression-free survival.

modifications that affect tumor cell adhesion, signaling, invasion and immune surveillance [23-25]. Interestingly, in our cohort, a high baseline level of MGAT4B expression was correlated with a favorable response and improved PFS. Spatial transcriptomics analysis in immunotherapy for NSCLC highlighted MGAT4B as a specific gene associated with therapeutic efficacy in lung squamous cell carcinoma, which may suggest the potential value of MGAT4B for

use as a predictive biomarker for selecting patients who might benefit from immunotherapy. We speculate that MGAT4B-mediated glycosylation enhances the presentation of tumor antigens or immune cell infiltration, which may sensitize the tumor to immune checkpoint blockade.

Our composite risk score based on these three proteins achieved superb predictive accuracy (AUC = 0.946) (i.e. better than any single protein), but more importantly, it could stratify patients into a group of high and low-risk patients with dramatically different PFS and OS. These findings suggest that this risk score could act as a non-invasive tool to assist in clinical decisions, for example, by identifying the patients who are most likely to have pathological response after neo adjuvant immunochemotherapy. Compared with existing markers such as PD-L1 expression and tumor mutational burden (TMB), both of which are constrained by tumor heterogeneity and sample dependency, our serum model in its current form has great advantages in terms of being minimally invasive, dynamic and representative of the host's systemic immune state. Direct comparison of our results with those of PD-L1 was precluded by incomplete data in our cohort; we think, however, that our three-protein signature complements rather than competes with available biomarkers and may be useful in patient selection when used in combination.

Our study used DIA mass spectrometry to undertake quantitative analysis of 1,802 serum proteins, and thus this study had advantages over previous proteomics studies. By fully employing serum DIA proteomics, we elaborated proteomic characteristics associated with immune therapy responses in NSCLC. Our work promotes the big picture for the greater use of serum proteomics in biomarker research in immuno oncology.

Another limitation of this study is that our research is a retrospective study with a small cohort (sample number n = 72), leading to selection bias and risk of overfitting; although we did LASSO regression analysis and stratified Bootstrap internal validation, prospective validation in a larger cohort study still needs to be performed. Second, our study population included Chinese patients with resectable stage IIIA/IIIB NSCLC, which may not generalize the results to other races or advanced patients. Third, because of scarce tissue, the samples of 52.5% patients were without an expression of baseline PD-L1, so the comparison with our proteomics model cannot be made. Fourth, the mechanistic reason for the expression of SERPINE2, DAZAP1, MGAT4B and response to immunotherapy is presumptive and needs to be validated by functional experiments and single-cell transcriptome tests.

Conclusions

In summary, this study discovered a serum triprotein biomarker combination (SERPINE2, DAZAP1, MGAT4B) by analysis of DIA proteomics, which could predict the effectiveness of immunotherapy, useful for stratification of patient survival of NSCLC. The composite risk score model is a promising, noninvasive tool for screening immunotherapy beneficiaries. The results of this study show the application of serum proteomics for translational biomarker discovery, and provide the basis for studying the mechanism of interaction between circulating proteins and antitumor immune response.

Acknowledgements

The authors would like to thank the staff of Tangdu Hospital, Fourth Military Medical University, for their administrative and technical support during data collection and sample processing.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Tao Jiang, Department of Thoracic Surgery, Tangdu Hospital, Air Force Medical University, No. 569 Xinxi Road, Baqiao District, Xi'an 710038, Shaanxi, China. E-mail: taojiangpro@163.com

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021; 71: 209-249.

[2] Zeng H, Chen W, Zheng R, Zhang S, Ji JS, Zou X, Xia C, Sun K, Yang Z, Li H, Wang N, Han R, Liu S, Li H, Mu H, He Y, Xu Y, Fu Z, Zhou Y, Jiang J, Yang Y, Chen J, Wei K, Fan D, Wang J, Fu F, Zhao D, Song G, Chen J, Jiang C, Zhou X, Gu X, Jin F, Li Q, Li Y, Wu T, Yan C, Dong J, Hua Z, Baade P, Bray F, Jemal A, Yu XQ and He J. Changing cancer survival in China during 2003-15: a pooled analysis of 17 population-based cancer registries. Lancet Glob Health 2018; 6: e555-e567.

[3] Sorin M, Prosty C, Ghaleb L, Nie K, Katergi K, Shahzad MH, Dubé LR, Atallah A, Swaby A, Dankner M, Crump T, Walsh LA, Fiset PO, Sepesi B, Forde PM, Cascone T, Provencio M and Spicer JD. Neoadjuvant Chemoimmunotherapy for NSCLC: a systematic review and meta-analysis. JAMA Oncol 2024; 10: 621-633.

[4] Forde PM, Spicer J, Lu S, Provencio M, Mitsudomi T, Awad MM, Felip E, Broderick SR, Brahmer JR, Swanson SJ, Kerr K, Wang C, Ciuleanu TE, Saylors GB, Tanaka F, Ito H, Chen KN, Liberman M, Vokes EE, Taube JM, Dorange C, Cai J, Fiore J, Jarkowski A, Balli D, Sausen M, Pandya D, Calvet CY and Girard N; CheckMate 816 Investigators. Neoadjuvant nivolumab plus chemotherapy in resectable lung cancer. N Engl J Med 2022; 386: 1973-1985.

[5] Wakelee H, Liberman M, Kato T, Tsuboi M, Lee SH, Gao S, Chen KN, Dooms C, Majem M, Eigendorff E, Martinengo GL, Bylicki O, Rodríguez-Abreu D, Chaft JE, Novello S, Yang J, Keller SM, Samkari A and Spicer JD; KEYNOTE-671 Investigators. Perioperative pembrolizumab for early-stage non-small-cell lung cancer. N Engl J Med 2023; 389: 491-503.

[6] Lei J, Zhao J, Gong L, Ni Y, Zhou Y, Tian F, Liu H, Gu Z, Huang L, Lu Q, Wang X, Sun J, Yang E, Wang T, Zhong D, Wang J, Zhao Z, Liu Z, Wang C, Wang X, Lei G, Yan X and Jiang T. Neoadjuvant camrelizumab plus platinum-based chemotherapy vs chemotherapy alone for Chinese patients with resectable stage IIIA or IIIB (T3N2) non-small cell lung cancer: the TD-FOREKNOW randomized clinical trial. JAMA Oncol 2023; 9: 1348-1355.

[7] Gandhi L, Rodríguez-Abreu D, Gadgeel S, Esteban E, Felip E, De Angelis F, Domine M, Clingan P, Hochmair MJ, Powell SF, Cheng SY, Bischoff HG, Peled N, Grossi F, Jennens RR, Reck M, Hui R, Garon EB, Boyer M, Rubio-Viqueira B, Novello S, Kurata T, Gray JE, Vida J, Wei Z, Yang J, Raftopoulos H, Pietanza MC and Garassino MC; KEYNOTE-189 Investigators. Pembrolizumab plus chemotherapy in metastatic non-small-cell lung cancer. N Engl J Med 2018; 378: 2078-2092.

[8] West H, McCleod M, Hussein M, Morabito A, Rittmeyer A, Conter HJ, Kopp HG, Daniel D, McCune S, Mekhail T, Zer A, Reinmuth N, Sadiq A, Sandler A, Lin W, Ochi Lohmann T, Archer V, Wang L, Kowanetz M and Cappuzzo F. Atezolizumab in combination with carboplatin plus nab-paclitaxel chemotherapy compared with chemotherapy alone as first-line treatment for metastatic non-squamous non-small-cell lung cancer (IMpower130): a multicentre, randomised, open-label, phase 3 trial. Lancet Oncol 2019; 20: 924-937.

[9] Di Federico A, Alden SL, Smithy JW, Ricciuti B, Alessi JV, Wang X, Pecci F, Lamberti G, Gandhi MM, Vaz VR, Spurr LF, Sholl LM, Pfaff KL, Rodig SJ, Li YY, Cherniack AD, Nishino M, Johnson BE and Awad MM. Intrapatient variation in PD-L1 expression and tumor mutational burden and the impact on outcomes to immune checkpoint inhibitor therapy in patients with non-small-cell lung cancer. Ann Oncol 2024; 35: 902-913.

[10] John N, Schlintl V, Sassmann T, Lindenmann J, Fediuk M, Wurm R, Douschan P, Zacharias M, Kalson L, Posch F, Absenger G, Brcic L, Jost PJ and Terbuch A. Longitudinal analysis of PD-L1 expression in patients with relapsed NSCLC. J Immunother Cancer 2024; 12: e008592.

[11] Arnon J, Dinstag G, Tirosh O, Gugel L, Kinar Y, Gottlieb T, Elia A, Rottenberg Y, Nechushtan H, Tabi M, Blumenfeld P, Pikarsky E, Beker T, Aharonov R and Popovtzer A. Predictive value of ENLIGHT-DP in patients with metastatic lung adenocarcinoma treated with immune checkpoint inhibitors and platinum chemotherapy directly from histopathology slides using inferred transcriptomics. J Immunother Cancer 2025; 13: e010132.

[12] Kok VC and Yu CC. Cancer-derived exosomes: their role in cancer biology and biomarker development. Int J Nanomedicine 2020; 15: 8019-8036.

[13] Dasari A, Morris VK, Allegra CJ, Atreya C, Benson AB 3rd, Boland P, Chung K, Copur MS, Corcoran RB, Deming DA, Dwyer A, Diehn M, Eng C, George TJ, Gollub MJ, Goodwin RA, Hamilton SR, Hechtman JF, Hochster H, Hong TS, Innocenti F, Iqbal A, Jacobs SA, Kennecke HF, Lee JJ, Lieu CH, Lenz HJ, Lindwasser OW, Montagut C, Odisio B, Ou FS, Porter L, Raghav K, Schrag D, Scott AJ, Shi Q, Strickler JH, Venook A, Yaeger R, Yothers G, You YN, Zell JA and Kopetz S. ctDNA applications and integration in colorectal cancer: an NCI colon and rectal-anal task forces whitepaper. Nat Rev Clin Oncol 2020; 17: 757-770.

[14] Liu MC, Oxnard GR, Klein EA, Swanton C and Seiden MV; CCGA Consortium. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. Ann Oncol 2020; 31: 745-759.

[15] Zaporozhchenko IA, Ponomaryova AA, Rykova EY and Laktionov PP. The potential of circulating cell-free RNA as a cancer biomarker: challenges and opportunities. Expert Rev Mol Diagn 2018; 18: 133-145.
[16] Sun R, A J, Yu H, Wang Y, He M, Tan L, Cheng H, Zhang J, Wang Y, Sun X, Lyu M, Qu M, Huang L, Li Z, Zhang W, Ma K, Dong Z, Ge W, Zhang Y, Ding X, Yang B, Hou J, Xu C, Wang L, Zhu Y, Guo T, Gao X and Yang C. Proteomic landscape profiling of primary prostate cancer reveals a 16-protein panel for prognosis prediction. Cell Rep Med 2024; 5: 101679.
[17] Zhang J, Wu Q, Zhu L, Xie S, Tu L, Yang Y, Wu K, Zhao Y, Wang Y, Xu Y, Chen X, Ma S and Zhang S. SERPINE2/PN-1 regulates the DNA damage response and radioresistance by activating ATM in lung cancer. Cancer Lett 2022; 524: 268-283.
[18] Zhang S, Jia X, Dai H, Zhu X, Song W, Bian S, Wu H, Chen S, Tang Y, Chen J, Jin C, Zhou M, Xie H, Zheng S and Song P. SERPINE2 promotes liver cancer metastasis by inhibiting c-Cbl-mediated EGFR ubiquitination and degradation. Cancer Commun (Lond) 2024; 44: 384-407.
[19] Chen WJ, Dong KQ, Pan XW, Gan SS, Xu D, Chen JX, Chen WJ, Li WY, Wang YQ, Zhou W, Rini B and Cui XG. Single-cell RNA-seq integrated with multi-omics reveals SERPINE2 as a target for metastasis in advanced renal cell carcinoma. Cell Death Dis 2023; 14: 30.
[20] Qiu L, Li W, Zhang L, Zhang X, Zhao H, Miyagishi M, Wu S and Kasim V. p52-ZER6/DAZAP1 axis promotes ferroptosis resistance and colorectal cancer progression via regulating SLC7A11 mRNA stabilization. Acta Pharm Sin B 2025; 15: 2039-2058.
[21] Zhang P, Wang W, Xiang H, Zhou Y, Peng Q, Liu G, Xu ZX and Lu L. DAZAP1 maintains gastric cancer stemness by inducing mitophagy. JCI Insight 2025; 10: e175422.
[22] Wang X, Fan H, Ye X, Hu Y, Xiao Y, Zhang M, Xu Y, Song J and Luo Y. RNA-binding protein DAZAP1 accelerates the advancement of pancreatic cancer by inhibiting ferroptosis. Eur J Med Res 2025; 30: 3.
[23] Sharma B, Aswin K, Jain T, Nasreen A, Aggarwal A, Subramaniam YJ, Rengaraju J, Agrawal SJ, Bhatt M, Paul B, Chandrasekaran K, Yadav A, Soni J, Ujjainiya R, Akhter MQ, Pandey R, Suresh S, Sampathkumar SG and Natarajan VT. Mgat4b-mediated selective N-glycosylation regulates melanocyte development and melanoma progression. Proc Natl Acad Sci U S A 2025; 122: e2423831122.
[24] Zhang K, Xie H, Zhao F and Huang Y. Identification of the gene signatures related to NK/T cell communication to evaluate the tumor microenvironment and prognostic outcomes of patients with prostate adenocarcinoma. Front Immunol 2025; 16: 1564784.
[25] Maeda K and Taniguchi N. Enzyme assay of N-acetylglucosaminyltransferase-IV (GnT-IVa,b, MGAT4A,B). 2021 Oct 1 [updated 2022 Mar 23]. In: Nishihara S, Angata K, Aoki-Kinoshita KF, Hirabayashi J, editors. Glycoscience Protocols (GlycoPODv2) [Internet]. Saitama (JP): Japan Consortium for Glycobiology and Glycotechnology; 2021.