Asian Journal of Current Research in Clinical Cancer
ISSN: 3062-4444
2024, Volume 4, Issue 1, Page No: 79-89
Copyright CC BY-NC-SA 4.0
Available online at: www.galaxypub.co/page/journals

Galaxy Publication

Galaxy Publication

Machine Learning–Derived Intrinsic Gene Signatures Predict Outcomes to PD-1 Inhibitors in Lung Adenocarcinoma

S. O'Connell[1*], D. Murphy[1], F. Kelly[1]

[1]Department of Cancer Research, Faculty of Medicine, University of Dublin, Dublin, Ireland.

***E-mail** ✉ dublin.research.67@protonmail.com

Received: 12 October 2023; **Revised:** 29 January 2024; **Accepted:** 04 February 2024

ABSTRACT

To better understand why patients with non–small cell lung cancer (NSCLC) experience variable benefit from PD-1 inhibitors, we trained machine learning models on transcriptomic data from 57 treated individuals. When ranking the genes most strongly linked with therapeutic outcome, lung adenocarcinoma (LUAD) displayed a pronounced enrichment of tumor-intrinsic genes (69%), in contrast to both the broader NSCLC cohort (36%) and lung squamous cell carcinoma (LUSC) (33%). A signature constructed from LUAD-specific intrinsic transcripts provided the most reliable classification of treatment response, yielding a mean ROC AUC of 0.957 and an accuracy of 0.9—substantially higher than signatures derived from extrinsic programs or from intrinsic sets in NSCLC or LUSC. LUAD patients with elevated intrinsic-signature activity showed significantly longer overall survival ($p = 0.034$). Functional annotation of the LUAD intrinsic genes highlighted enrichment of pathways related to cell-cycle control and senescence. This signature also demonstrated positive associations with several immune checkpoint molecules, including CD274, LAG3, and PDCD1LG2 (Spearman's $\rho > 0.25$). The marked divergence in intrinsic transcriptional patterns between LUAD and LUSC suggests that intrinsic gene programs may serve as a particularly informative biomarker for predicting PD-1 inhibitor benefit in LUAD.

Keywords: PD-L1, Non-small cell lung cancer, PD-1, PD-1/PD-L1-targeted therapy, Immunotherapy, Biomarkers

How to Cite This Article: O'Connell S, Murphy D, Kelly F. Machine Learning–Derived Intrinsic Gene Signatures Predict Outcomes to PD-1 Inhibitors in Lung Adenocarcinoma. Asian J Curr Res Clin Cancer. 2024;4(1):79-89. https://doi.org/10.51847/MBFVGm4SXy

Introduction

The introduction of programmed cell death protein-1 (PD-1) inhibitors into the therapeutic landscape of advanced non–small cell lung cancer (NSCLC) has transformed clinical management [1]. Although PD-L1 immunohistochemistry is routinely used to estimate the likelihood of benefit, it is clear that PD-L1 expression alone does not fully account for treatment outcomes, as responses can occur even in patients whose tumors show minimal staining [2]. Numerous additional biomarkers—including tumor mutational burden, microbial composition of the gut, and specific genetic alterations such as *STK11* or *β-catenin* mutations—have been explored to improve prediction, but their clinical reliability remains uncertain [3].

Attention has increasingly turned to transcriptomic information as a complementary predictor of immunotherapy efficacy. Early work by Ayers and collaborators demonstrated that a transcriptomic pattern linked to interferon-γ could stratify patient response across several tumor types, including melanoma, head and neck squamous cell carcinoma (HNSCC), and gastric cancer [4]. Independent studies later showed similar utility of this gene program in other HNSCC cohorts [5] and in NSCLC [6]. Because immune-related cytokines, such as interferon-γ and TNF-α, play key roles in regulating PD-1/PD-L1 biology, most prior efforts have focused almost exclusively on immune-driven (extrinsic) transcriptional signatures.

Recent findings, however, indicate that non-immune, tumor-derived signals also modulate PD-L1 and PD-1 expression. Intrinsic pathways influenced by oncogenic events, epigenetic remodeling, transcription factor

© 2024 Asian Journal of Current Research in Clinical Cancer

activity, and stemness-associated programs can shape sensitivity or resistance to PD-1 blockade [7]. Despite this, the transcriptional landscape associated with such intrinsic mechanisms has received little systematic investigation. The need to examine these pathways separately for distinct NSCLC subtypes is clear, given that lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) harbor fundamentally different mutational compositions that influence disease behavior and treatment options [8].

Machine learning provides a practical approach for dissecting complex biological data and is increasingly used in biomarker discovery [9]. Its application in oncology has improved prognostic modeling accuracy by 15–20% in recent evaluations [10]. Feature selection and classification serve as the backbone of most machine learning pipelines. In the current work, we implemented six classification strategies and used an analysis-of-variance (ANOVA)–based approach to identify relevant features.

To investigate the contribution of tumor-intrinsic and immune-associated transcriptional signals to PD-1 inhibitor therapy, we evaluated expression data from 57 NSCLC patients treated with PD-1 monotherapy. A total of 780 genes were categorized according to whether they reflect intrinsic or extrinsic biological processes. We assessed the predictive value of these transcriptional patterns, performed subtype-specific analyses for LUAD and LUSC, and explored enriched pathways among the most informative genes.

Materials and Methods

Patient cohort and PD-L1 analysis

The study was reviewed and approved by the Institutional Review Board of Ajou University School of Medicine (AJIRB-BMR-KSP-20-396), with informed consent waived because of the retrospective study design. Clinical records from 2016 to 2021 were screened to identify 57 patients with advanced NSCLC who received PD-1 inhibitors. Tissue samples were obtained either through biopsy or surgical resection. Patients who were not surgical candidates received PD-1 inhibitors as primary therapy, while those who underwent resection were treated upon recurrence or metastasis. The cohort comprised 31 LUAD and 26 LUSC cases.

Therapeutic responses were categorized as either responders (complete or partial response) or non-responders (stable disease or progression) following established criteria [11]. PD-L1 immunohistochemical staining was performed using the SP263 rabbit monoclonal antibody (Roche, Basel, Switzerland) and developed using the OptiView DAB Detection Kit on the Ventana BenchMark ULTRA system.

Transcriptomic profiling

Gene-expression profiling was conducted with the nCounter® Tumor Signaling 360 panel, which surveys a broad range of genes implicated in tumor biology and immune regulation [12]. The original panel includes 760 genes; we supplemented it with 20 additional genes known to participate in PD-L1 regulatory pathways, yielding a final set of 780 genes.

NanoString classifies genes in the panel into ten functional categories based on the conceptual framework proposed by Hanahan *et al*. [13]. We assigned the added genes to these same categories. Of the ten categories, those associated with immune suppression or tumor-associated inflammation were considered extrinsic; all others were grouped as intrinsic.

For our analyses of tumor-intrinsic biology, only intratumoral tissue was microdissected from formalin-fixed paraffin-embedded (FFPE) blocks. RNA extraction and hybridization with capture and reporter probes were carried out according to high-sensitivity procedures recommended by the manufacturer. Expression counts were obtained on the nCounter Digital Analyzer and normalized using the geometric mean of housekeeping genes and internal positive controls.

Machine learning approach and statistical analysis

To explore whether transcriptional patterns could predict clinical response to PD-1 blockade, we applied machine learning techniques to the mRNA data generated using the NanoString platform. Because the original dataset contained 780 genes, an initial reduction step was required to isolate the most informative features. We used analysis of variance (ANOVA)—a method commonly used in transcriptomic research—to identify genes whose expression differed significantly between responders and non-responders [14–16].

After ranking genes by their ANOVA scores, we next determined how many of these candidates were optimal for building predictive models. Gene subsets of varying sizes were evaluated, and the number that achieved the highest area under the receiver operating characteristic curve (AUC) was selected. Six supervised learning models

were used to compute AUC values for each gene subset: naïve Bayes (NB), neural networks (NN), random forests (RF), logistic regression (LR), support vector machines (SVM), and k-nearest neighbors (kNN). The AUC metric reflects the discriminatory capacity of the selected genes in distinguishing between treatment responders and non-responders.

Because the cohort size was limited, model performance was assessed using leave-one-out cross-validation (LOOCV). In this strategy, each sample is held out once as an independent test case, while all remaining samples are used for training, resulting in n training–testing cycles for n samples. For each model, we quantified performance using five criteria: AUC, accuracy, precision, recall (sensitivity), and F1 score (the harmonic mean of precision and recall). The model with the highest combined score across these five metrics was considered the best-performing approach. Predictions derived from this model were subsequently used for survival analyses.

Statistical analysis

Fold-change values, normalized expression levels, and associated p-values were obtained through the nSolver software suite (NanoString Technologies). Other statistical analyses were performed using IBM SPSS Statistics 25 (IBM, Armonk, NY, USA) and R version 3.5.3, with statistical significance defined as $p < 0.05$. Machine learning procedures were carried out using the Orange 3.27 software environment (Bioinformatics Laboratory, University of Ljubljana) [17]. Gene Set Enrichment Analysis (GSEA) was performed using version 4.0.3 of the GSEA software package [18], and pathway analysis based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) was conducted using DAVID Bioinformatics Resources 6.8 [19].

Results and Discussion

Feature selection and prediction modeling

Using ANOVA, we identified 100 genes whose expression showed the strongest association with response to PD-1 inhibition. Based on NanoString pathway annotations, these genes were separated into intrinsic and extrinsic categories. Among the 57 NSCLC cases, 36 genes belonged to intrinsic pathways and 64 to extrinsic pathways. In LUAD, 69 intrinsic and 31 extrinsic genes were identified, whereas in LUSC, the balance shifted to 33 intrinsic and 67 extrinsic genes.

To establish the optimal number of genes for predictive modeling, we compared the AUC values generated by the six machine learning algorithms for gene subsets ranging from 5 to 30 genes. Across the full NSCLC cohort, intrinsic pathways achieved their highest combined AUC score with 30 genes, while extrinsic pathways performed best with 5 genes. In LUAD, both intrinsic and extrinsic pathways reached peak predictive performance with 15 genes. In contrast, LUSC showed its highest AUC sum with 5 intrinsic genes and 10 extrinsic genes. The resulting gene sets for each pathway are listed in **Table 1**.

Table 1. The gene list of each pathway.

30 intrinsic pathway related genes in a total NSCLC
GREM1, AURKA, AXL, RRM2, CDC25A, PRIM1, CCNB1, ATP7A, PCLAF, LMNB1, ERBB2, NSD1, KAT6B, CDCA8, KIF2C, SRM, GPX1, PIK3CB, MYBL2, RUVBL1, UBE2C, RBBP5, FASLG, ARID4B, NCAPG, PIK3R5, CAV1, BAX, MYC, ITGA5.
5 extrinsic pathway related genes in a total NSCLC
IL15RA, CCR1, CCL2, CYBB, FCER1G
15 intrinsic pathway related genes in lung adenocarcinoma
MYC, RRM2, MAPK14, MYBL2, AXL, CDCA5, AKT1S1, TGFB1, CDK12, PCLAF, BAX, HMOX1, ARID4A, SLC1A5, ARID4B.
15 extrinsic pathway related genes in lung adenocarcinoma
CD209, CCL13, LAG3, BCL2L1, HLA-C, IL15RA, MSLN, TAP1, CD8A, NKG7, TAP2, MST1R, CCL5, CCL2, CYBB.
5 intrinsic pathway related genes in lung squamous cell carcinoma
SERPINE1, PIK3R5, PIK3CB, RPA3, KIF2C
10 extrinsic pathway related genes in lung squamous cell carcinoma
CCR1, FCER1G, CD38, GNLY, CYBB, IL15RA, CCL2, VAV1, CD274, GZMB

Across the 57 patients with NSCLC, we examined how well the intrinsic-pathway gene panel could distinguish responders using six independent learning algorithms. Although the individual AUC values differed by method, they collectively fell within a moderate–high performance window (0.71–0.87), and the combined contribution of the six methods resulted in an aggregate score of 4.835 **(Figure 1a)**. When the intrinsic signature was fed into the neural-network classifier, the resulting patient stratification produced clearly separated survival curves, reaching statistical significance (p = 0.037) **(Figure 1b)**. This suggests that the intrinsic component of the transcriptomic profile carries prognostic weight.

For the extrinsic-pathway gene subset—which was smaller and consisted of only five markers—the classification output again varied by algorithm but remained in a similar performance band (AUCs between 0.74 and 0.823). Summing the AUCs yielded 4.756 **(Figure 1c)**. Among the methods tested, the SVM-based stratification appeared to distinguish longer-surviving individuals most effectively, although the difference did not reach significance (p = 0.062) **(Figure 1d)**.

In contrast to these transcriptomic-based approaches, PD-L1 immunohistochemistry contributed essentially no predictive value. Its AUC remained close to the level of random classification (0.538) **(Figure 1e)**, and survival outcomes were indistinguishable between PD-L1 expression groups (p = 0.92) **(Figure 1f)**.

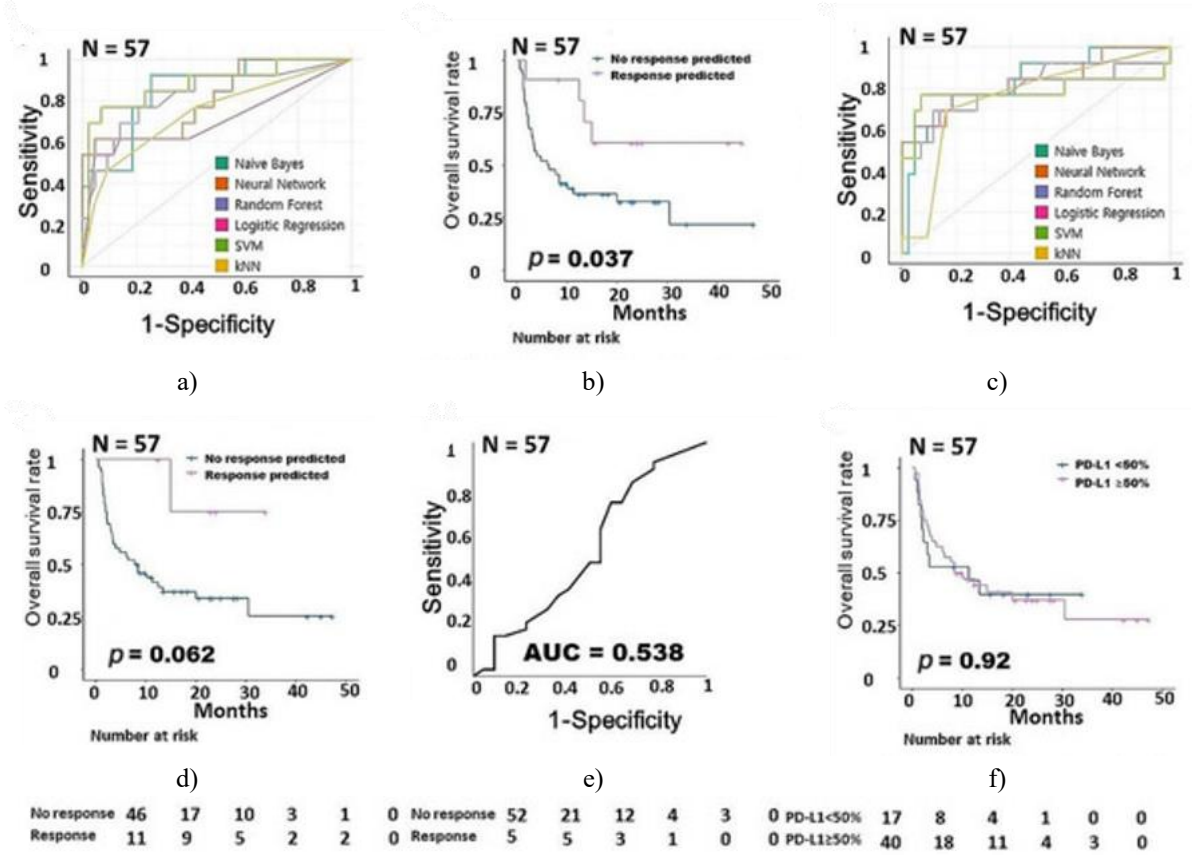

Figure 1. Modeling PD-1 inhibitor responsiveness in NSCLC. (a) AUCs for the intrinsic gene panel across six machine-learning approaches. (b) Kaplan–Meier survival curves based on the intrinsic signature. (c) AUCs for the extrinsic gene panel across six classifiers. (d) Survival curves according to the extrinsic signature. (e) AUC for PD-L1 expression. (f) Survival analysis based on PD-L1 status. Abbreviations: AUC, area under the ROC curve; NSCLC, non–small cell lung cancer; PD-1, programmed cell death protein 1; PD-L1, programmed cell death-ligand 1.

In LUAD patients, a set of 15 genes associated with intrinsic pathways demonstrated consistently high predictive strength, with AUC values exceeding 0.9 for every classifier **(Figure 2a)**. Such high AUCs indicate very strong discriminatory power, highlighting the ability of these intrinsic transcripts to differentiate likely responders from non-responders. When survival outcomes were analyzed using predictions from the logistic regression model, individuals classified as high-scoring according to the intrinsic signature showed significantly improved overall survival compared to the remaining patients (p = 0.034) **(Figure 2b).**

The corresponding extrinsic gene panel exhibited weaker performance, with lower AUCs across all six models **(Figure 2c)**. While the support vector machine identified a subset with somewhat better survival, the difference did not reach statistical significance (p = 0.071) **(Figure 2d)**. Immunohistochemical assessment of PD-L1 provided modest predictive information, with an AUC of 0.67 **(Figure 2e)**, and survival analysis confirmed that PD-L1 expression alone was not associated with differences in prognosis (p = 0.91) **(Figure 2f)**.

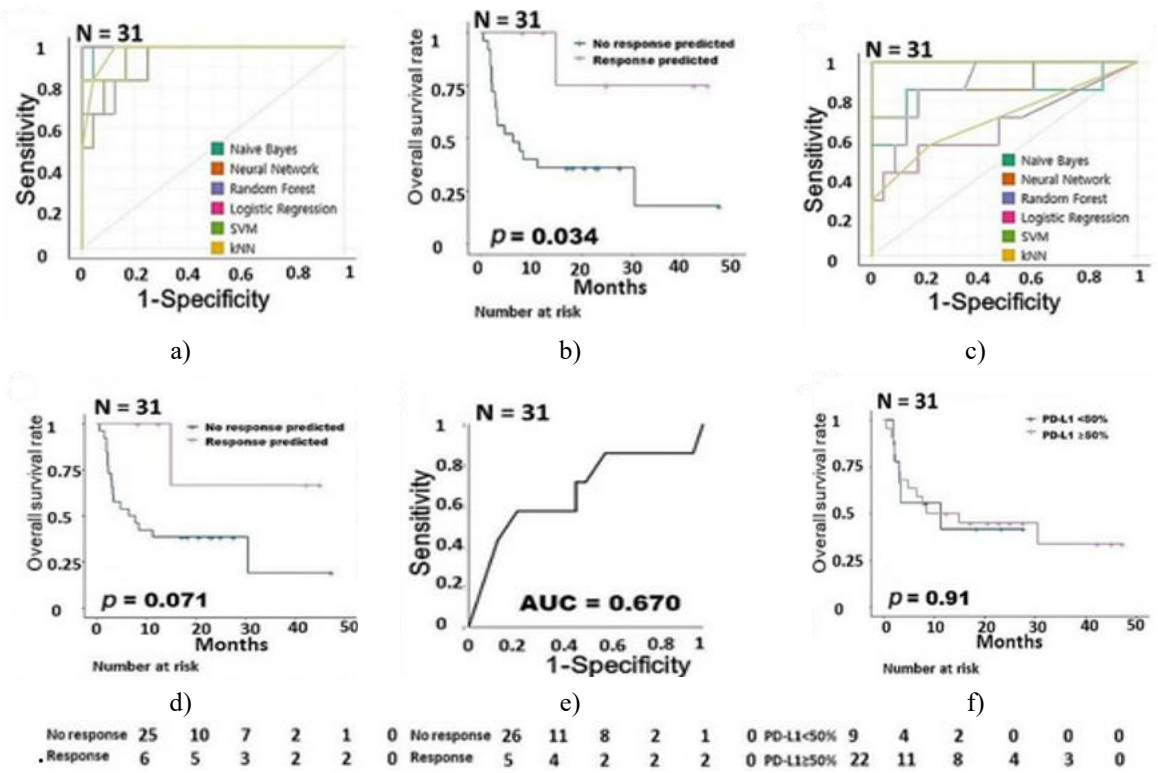

Figure 2.

Figure 2 To investigate how transcriptional programs relate to PD-1 inhibitor response, we analyzed intrinsic and extrinsic gene panels separately for LUAD and LUSC. In LUAD, a 15-gene intrinsic signature demonstrated exceptionally high predictive power, with all machine-learning algorithms achieving AUCs above 0.9. This strong signal allowed logistic regression–based predictions to identify a patient subgroup with significantly improved survival outcomes (p = 0.034), highlighting the clinical relevance of intrinsic tumor programs. In contrast, the extrinsic gene panel in LUAD showed weaker predictive performance, and although the SVM model identified a subgroup with longer survival, the difference did not reach statistical significance (p = 0.071). Notably, PD-L1 expression in this cohort provided only modest predictive value (AUC = 0.67) and did not correlate with survival (p = 0.91).

The LUSC cohort exhibited a different pattern. A smaller set of five intrinsic genes yielded moderate predictive ability, with AUCs ranging from approximately 0.71 to 0.90. Neural network–based predictions using these genes successfully stratified patients, identifying a subgroup with significantly better survival (p = 0.0085). The 10-gene extrinsic panel in LUSC showed comparable overall predictive performance, and the SVM classifier separated a group with improved prognosis (p = 0.018). Similar to LUAD, PD-L1 expression in LUSC offered limited predictive power (AUC = 0.67) and did not influence survival outcomes (p = 0.98).

Collectively, these results indicate that intrinsic transcriptional programs, rather than PD-L1 expression, are the most robust indicators of response to PD-1 inhibitors in NSCLC, with notable differences between LUAD and LUSC.

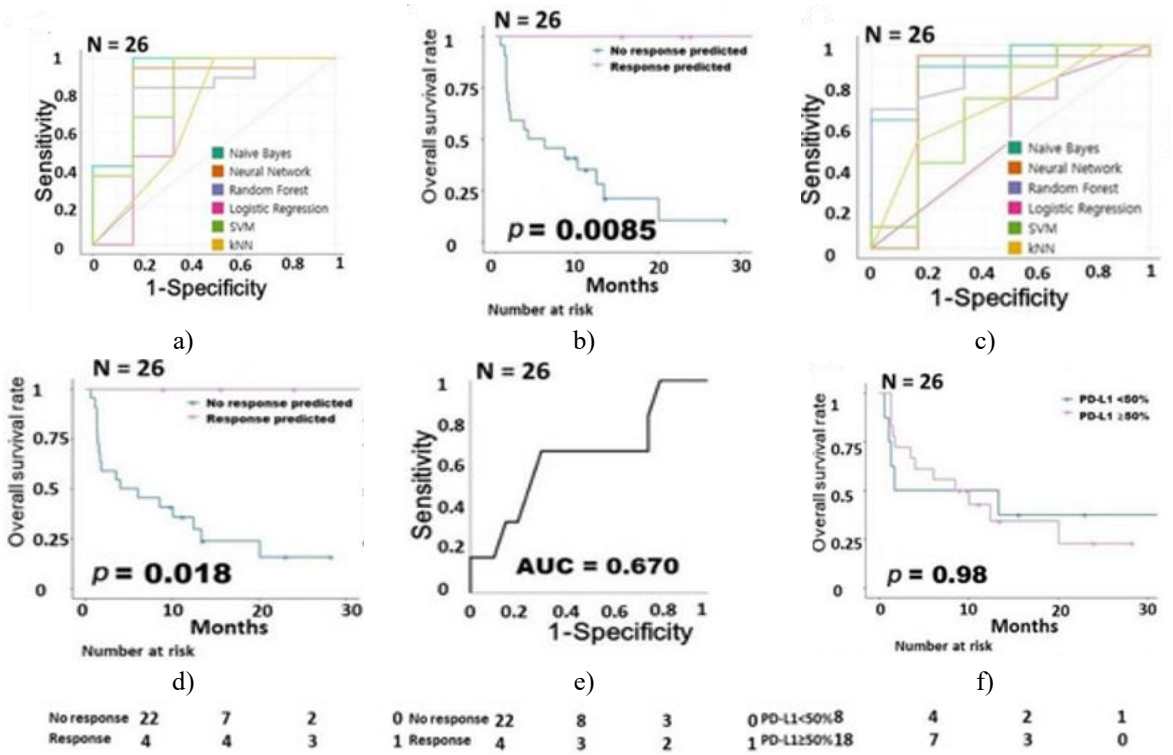

Figure 3.

Figure 3. Building on previous studies, the IFN-γ–associated gene signature described by Ayers *et al*. remains one of the most widely cited immune-related transcriptional predictors of PD-1 inhibitor response [4]. Their original panel included 18 genes; in our study. When applied to LUAD, the naïve Bayes classifier achieved an AUC of 0.714, while the other five machine-learning methods performed less effectively, with AUCs below 0.6. Moreover, survival analysis using logistic regression did not show a significant correlation between this immune signature and overall survival in LUAD ($p = 0.43$). In contrast, for LUSC, logistic regression produced the highest AUC at 0.85, and additional classifiers (NB, NN, RF) exceeded 0.7. Notably, in LUSC, the immune signature predicted improved prognosis in the logistic regression survival analysis ($p = 0.0019$), highlighting a subtype-specific effect.

Given the known impact of EGFR mutations on immunotherapy response, we conducted a subgroup analysis restricted to EGFR-negative patients, as all EGFR-mutant cases in our cohort were non-responders. In this subgroup, the 15-gene intrinsic signature achieved AUCs between 0.87 and 1 across classifiers (NB: 1; NN: 1; RF: 0.870; LR: 0.987; SVM: 0.987; kNN: 0.883; total AUC sum: 5.727), while the 15-gene extrinsic panel showed more variable performance (AUC range 0.5–1, total AUC sum: 4.98). These results indicate that intrinsic pathway genes maintain slightly higher predictive power even in EGFR-negative patients.

Survival analysis using TCGA data

To examine the broader relevance of intrinsic gene signatures beyond our PD-1 inhibitor–treated cohort, we leveraged TCGA datasets of patients who had not received immunotherapy. Intrinsic signature scores were calculated as the sum of normalized upregulated gene expression minus the sum of downregulated genes. Interestingly, higher intrinsic scores correlated with poorer survival in both LUAD ($p < 0.001$) and LUSC ($p = 0.022$), suggesting that these tumor-intrinsic programs may influence prognosis independently of PD-1 therapy.

Pathway enrichment analysis

To explore the biological functions associated with intrinsic and extrinsic genes, we performed KEGG pathway enrichment on genes identified via ANOVA. Upregulated and downregulated genes in responders were analyzed separately, and the top five pathways based on statistical significance were selected.

In LUAD, intrinsic genes upregulated in responders were predominantly associated with cell cycle regulation and cellular senescence, whereas downregulated intrinsic genes were enriched in MAPK signaling, focal adhesion,

and lysine degradation **(Figure 4a)**. Extrinsic genes upregulated in LUAD were entirely linked to immune-related processes, including cytokine–cytokine receptor interactions **(Figure 4b)**.

In LUSC, upregulated intrinsic genes mapped to cellular senescence and HIF-1 signaling pathways, while downregulated intrinsic genes were enriched in proteoglycans in cancer and Hedgehog signaling **(Figure 4c)**. Upregulated extrinsic genes in LUSC were similarly immune-focused, including pathways such as natural killer cell–mediated cytotoxicity **(Figure 4d)**. KEGG analysis was not performed for downregulated extrinsic genes in either subtype due to insufficient gene numbers.

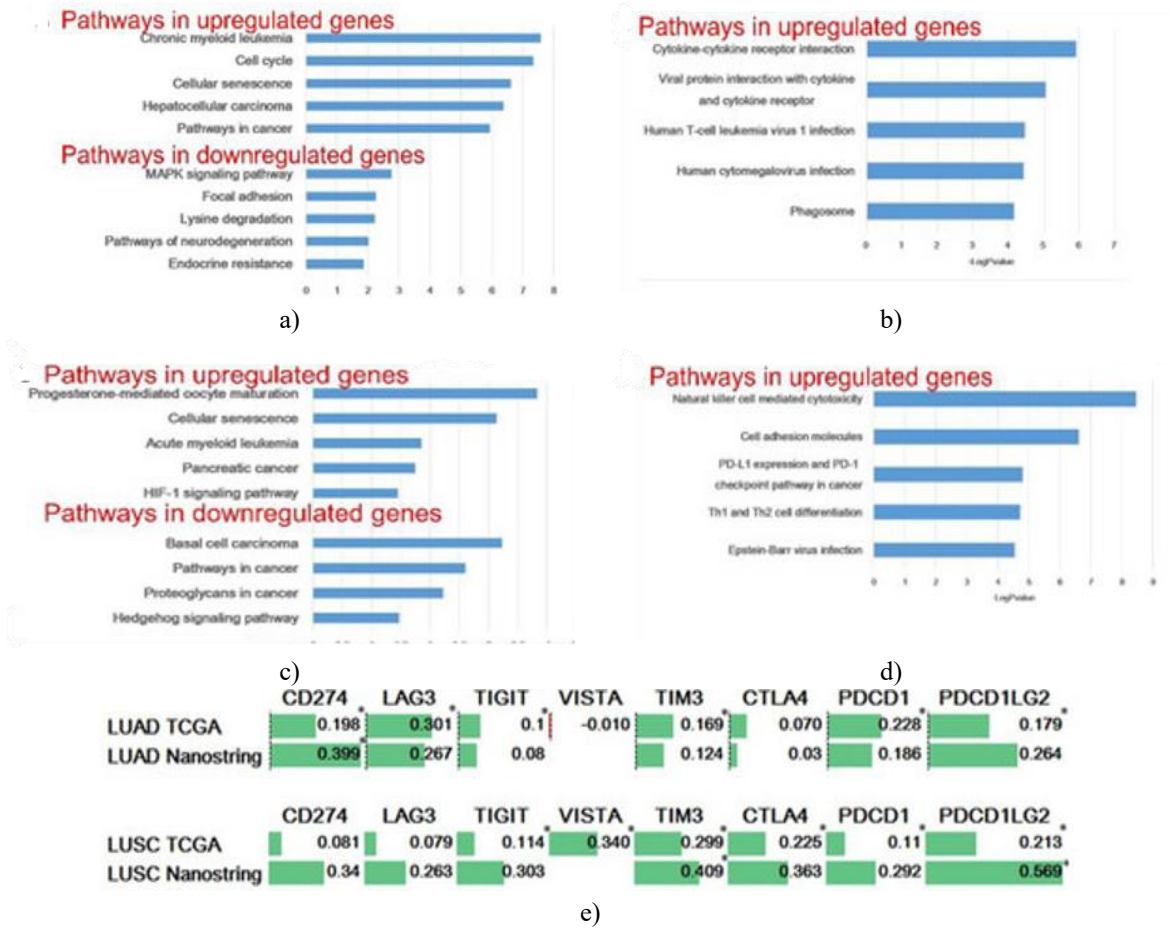

Figure 4. Illustrates the comparative findings of ^{68}Ga-PSMA and ^{68}Ga-DOTA-RM2 PET/MRI. Panel (a) shows the overall patient-based detection rates for the two radiotracers. Panels (b), (c), and (d) display detection rates stratified according to serum PSA levels at the time of imaging, PSA doubling time, and Gleason score prior to radical treatment, respectively. Statistically significant differences between the tracers are indicated by an asterisk (*), representing $p < 0.05$.

Figure 4 To explore whether high intrinsic gene expression reflects activation of immune checkpoint pathways, we analyzed correlations between the intrinsic signature and multiple checkpoint targets in both the NanoString and TCGA datasets. In LUAD samples from TCGA, the intrinsic signature showed positive associations with CD274, LAG3, TIGIT, TIM3, PDCD1, and PDCD1LG2, with Spearman correlation coefficients ranging from 0.10 to 0.30. Similarly, in the NanoString LUAD cohort, the signature correlated positively with CD274, LAG3, and PDCD1LG2, all exceeding 0.25. In LUSC, TCGA data revealed significant positive correlations of the intrinsic signature with TIGIT, VISTA, TIM3, CTLA-4, PDCD1, and PDCD1LG2 (coefficients 0.11–0.34), while NanoString measurements showed consistent positive correlations with all seven checkpoint targets evaluated, with coefficients above 0.25. These findings suggest that tumors with high intrinsic gene expression may engage multiple immune-inhibitory pathways.

Next, we compared transcriptional profiles and predictive performance across NSCLC, LUAD, and LUSC. Averaging AUCs and accuracy values from six machine-learning models revealed that the intrinsic signature in

LUAD was the most powerful predictor of PD-1 inhibitor response (AUC = 0.957, accuracy = 0.9), outperforming both intrinsic and extrinsic signatures in the overall NSCLC cohort (AUC = 0.805 and 0.792, accuracy = 0.816 and 0.81, respectively). In contrast, LUSC intrinsic signatures demonstrated moderate predictive capacity (AUC = 0.797, accuracy = 0.833), comparable to NSCLC overall.

Examination of gene composition revealed that LUAD's top 100 genes from ANOVA were predominantly intrinsic (69%), whereas NSCLC and LUSC were dominated by extrinsic genes (64% and 67%, respectively). This indicates that LUAD and LUSC may be governed by distinct PD-L1/PD-1 regulatory mechanisms. Assessment of overlapping genes with the overall NSCLC panel showed greater concordance in LUSC than LUAD (56 vs. 41 genes). In LUAD, overlap was balanced between intrinsic and extrinsic genes (21 vs. 20), whereas in LUSC, extrinsic genes predominated among overlapping genes (44 vs. 12). Across both subtypes, only 15 genes were shared, of which only two (GREM1, AURKA) were intrinsic, while the remaining 13 (including LAG3, IL15RA, TAP1, NKG7, CCL2, CYBB, CD68, CD274, CD14, CCR1, FCGR3A/B, STAT1, PDCD1LG2) were classified as extrinsic.

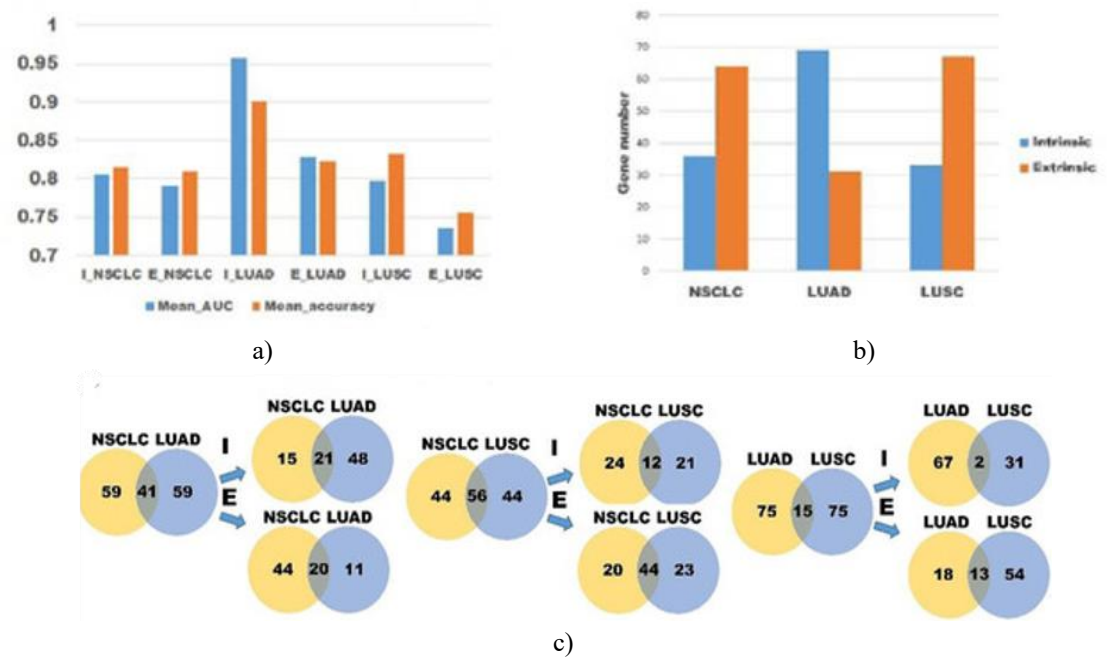

Figure 5. Differences in transcriptional patterns among NSCLC, LUAD, and LUSC.

Discussion

Our pathway enrichment analysis highlighted that intrinsic genes in LUAD were significantly associated with cell cycle regulation, while both LUAD and LUSC shared enrichment in the cellular senescence pathway. These findings align with prior evidence linking PD-L1 expression to cell cycle–related mechanisms. For instance, inhibition of CDK4/6 or depletion of cyclin D1 has been shown to elevate PD-L1 protein levels, and PD-L1 expression varies across cell cycle phases in cancer cell lines [20, 21]. NEK2, a key regulator of centrosome separation and microtubule stability, can interact with PD-L1 and prevent its ubiquitin-mediated degradation, and dual NEK2/PD-L1 inhibition has demonstrated potent anti-tumor effects in preclinical models [22].

The cellular senescence pathway is also closely connected to PD-L1/PD-1 regulation. Senescent stromal cells secrete amphiregulin, which enhances PD-L1 expression and promotes an immunosuppressive tumor microenvironment, while amphiregulin blockade can inhibit tumor progression and reduce chemoresistance [23]. Similarly, therapy-induced senescence and the associated secretory phenotype (SASP) have been shown to improve responses to PD-1 blockade in ovarian cancer and melanoma by recruiting CD8+ T cells and overcoming checkpoint resistance [24–26].

Within LUAD, several intrinsic genes in our signature have established links to immune checkpoint regulation. MYC directly induces PD-L1 transcription, while inhibition of RRM2 reduces PD-L1 levels and enhances tumor regression when combined with PD-1 blockade [27, 28]. MYBL2 correlates positively with multiple checkpoints,

including TIGIT, PDCD1, and LAG3 [29], and AXL and CDCA5 expression are similarly associated with PD-L1/PD-1 axis activation [30, 31].

At the cohort level, intrinsic genes dominate the top ANOVA-ranked genes in LUAD, whereas extrinsic genes are more prevalent in NSCLC overall due to shared extrinsic programs between LUAD and LUSC. This distinction underscores the differential regulation of PD-L1/PD-1 pathways between subtypes. Supporting this, Ayers' extended immune signature predicted prognosis in LUSC but not LUAD, further indicating that intrinsic transcriptional programs are more informative for LUAD patients.

Interestingly, while high intrinsic gene signatures predicted favorable outcomes in our PD-1–treated cohort, TCGA patients who had not received checkpoint inhibitors showed worse survival in the high-signature group. This suggests that tumors with elevated intrinsic signatures may harbor strong immune-inhibitory signals, leading to poor prognosis in the absence of therapy, but are more likely to benefit from PD-1 blockade. Thus, identifying patients with high intrinsic signatures could guide the selection of immunotherapy.

Moreover, several intrinsic genes in LUAD represent potential targets for combination therapy. FDI-6, a forkhead domain inhibitor, suppresses MYBL2 and FOXM1 activities and inhibits LUAD cell proliferation [32], while the AXL inhibitor R428 has shown survival benefits in breast cancer [33]. Early clinical trials are already exploring co-inhibition strategies, such as PD-1/PD-L1 blockade with TGF-β inhibitors or RRM2 suppression, to enhance CD8+ T cell–mediated anti-tumor responses [28, 34]. These observations point to the potential for precision combination therapies in LUAD guided by intrinsic gene signatures.

Study limitations

Several factors limit the interpretation of our findings. First, the sample size in this study was modest, and we were unable to confirm the results using an independent cohort. While many previous studies on PD-1 inhibitors have emphasized genomic alterations, research examining transcriptional profiles is still limited. Moreover, most of these studies have focused on immune-related genes without performing detailed analyses by tumor cell subtype, which complicates direct external validation. Second, our designation of intrinsic versus extrinsic pathways relied on the pathway annotations provided by NanoString. Because some genes may participate in multiple signaling pathways, it is possible that certain genes categorized as intrinsic could also influence extrinsic mechanisms.

Conclusion

This study highlights novel intrinsic gene signatures in LUAD and LUSC that provide valuable predictive information regarding responsiveness to PD-1 inhibitor therapy. The intrinsic signature appears especially relevant in LUAD, suggesting that tumor-intrinsic mechanisms play a prominent role in modulating the PD-L1/PD-1 axis in this subtype. Importantly, patients with high intrinsic gene expression may have a poor prognosis if untreated, indicating that these individuals could benefit substantially from PD-1 inhibitor therapy. These findings support the potential use of intrinsic transcriptional profiling to guide personalized immunotherapy strategies.

Acknowledgments: None

Conflict of Interest: None

Financial Support: This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT (NRF-2020R1A2C1100568 for Young Wha Koh). The funding provider had no role in research design, data collection and analysis, publication decisions, or manuscript preparation.

Ethics Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of the Ajou University School of Medicine (AJIRB-BMR-KSP-20-396). Informed consent was waived due to the retrospective study design.

References

1. Garon EB, Rizvi NA, Hui R, Leighl N, Balmanoukian AS, Eder JP, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. N Engl J Med. 2015;372(21):2018–28.
2. Dempke WCM, Fenchel K, Dale SP. Programmed cell death ligand-1 (PD-L1) as a biomarker for non-small cell lung cancer treatment. Transl Lung Cancer Res. 2018;7(Suppl 2):S275–9.
3. Havel JJ, Chowell D, Chan TA. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. Nat Rev Cancer. 2019;19(3):133–50.
4. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. J Clin Invest. 2017;127(8):2930–40.
5. Haddad RI, Seiwert TY, Chow LQ, Gupta S, Weiss J, Gluck I, et al. Influence of tumor mutational burden and PD-L1 expression on response to pembrolizumab. J Immunother Cancer. 2022;10(2):e003026.
6. Damotte D, Warren S, Arrondeau J, Boudou-Rouquette P, Mansuet-Lupo A, Biton J, et al. Tumor inflammation signature and anti-PD-1 benefit. J Transl Med. 2019;17(1):357.
7. Hudson K, Cross N, Jordan-Mahy N, Leyland R. Roles of PD-L1 and PD-1 in immunotherapy. Front Immunol. 2020;11:568931.
8. Meng F, Zhang L, Ren Y, Ma Q. Genomic alterations of lung adenocarcinoma vs squamous carcinoma. J Cell Physiol. 2019;234(7):10918–25.
9. Mamoshina P, Volosnikova M, Ozerov IV, Putin E, Skibina E, Cortese F, et al. Machine learning on human muscle transcriptomics. Front Genet. 2018;9:242.
10. Cruz JA, Wishart DS. Applications of machine learning in cancer prediction. Cancer Inform. 2007;2:59–77.
11. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. Revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228–47.
12. Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, et al. Direct multiplexed gene expression measurement. Nat Biotechnol. 2008;26(3):317–25.
13. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. Cell. 2011;144(5):646–74.
14. Cui S, Wu Q, West J, Bai J. Low-expression genes and PAH disease. PLoS Comput Biol. 2019;15(4):e1007264.
15. Shamsaei B, Gao C. Comparison of learning and statistical algorithms. In: IEEE-EMBS BHI Conference; 2016 Feb 24-27; Las Vegas. p. 296–9.
16. Abdulsalam S, Mohammed A, Ajao J, Babatunde R, Ogundokun R, Christopher C, et al. Evaluation of ANOVA and RFE with SVM on microarray data. In: European–Mediterranean–Middle Eastern Conf on Information Systems. Cham: Springer; 2020.
17. Demšar J, Curk T, Erjavec A, Gorup Č, Hočevar T, Milutinovič M, et al. Orange: Data mining toolbox in Python. J Mach Learn Res. 2013;14:2349–53.
18. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis. Proc Natl Acad Sci USA. 2005;102(43):15545–50.
19. Huang DW, Sherman BT, Lempicki RA. DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44–57.
20. Zhang J, Bu X, Wang H, Zhu Y, Geng Y, Nihira NT, et al. Cyclin D-CDK4 destabilizes PD-L1. Nature. 2018;553(7689):91–5.
21. Schulz D, Wetzel M, Eichberger J, Piendl G, Brockhoff G, Wege AK, et al. PD-L1 expression during cell cycle. Int J Mol Sci. 2021;22(23):13087.
22. Zhang X, Huang X, Xu J, Li E, Lao M, Tang T, et al. NEK2 inhibition triggers anti-pancreatic cancer immunity. Nat Commun. 2021;12:4536.
23. Xu Q, Long Q, Zhu D, Fu D, Zhang B, Han L, et al. Targeting AREG reduces PD-L1-mediated immunosuppression. Aging Cell. 2019;18(6):e13027.
24. Hao X, Zhao B, Zhou W, Liu H, Fukumoto T, Gabrilovich D, et al. Boosting SASP sensitizes ovarian tumors to checkpoint blockade. iScience. 2021;24:102016.
25. Jerby-Arnon L, Shah P, Cuoco MS, Rodman C, Su MJ, Melms JC, et al. A cancer cell program promoting T-cell exclusion. Cell. 2018;175(4):984–97.e924.
26. Ruscetti M, Morris JPT, Mezzadra R, Russell J, Leibold J, Romesser PB, et al. Senescence-induced vascular remodeling. Cell. 2020;181(2):424–41.e421.

27. Casey SC, Tong L, Li Y, Do R, Walz S, Fitzgerald KN, et al. MYC regulates immune response via CD47 and PD-L1. Science. 2016;352(6282):227–31.
28. Xiong W, Zhang B, Yu H, Zhu L, Yi L, Jin X. RRM2 regulates sensitivity to sunitinib and PD-1 blockade. Adv Sci. 2021;8(21):e2100881.
29. Chen X, Lu Y, Yu H, Du K, Zhang Y, Nan Y, et al. MYBL2 and pan-cancer prognosis. Cell Cycle. 2021;20(20):2291–308.
30. Tsukita Y, Fujino N, Miyauchi E, Saito R, Fujishima F, Itakura K, et al. Axl kinase drives immune checkpoint signaling. Mol Cancer. 2019;18:24.
31. Shen W, Tong D, Chen J, Li H, Hu Z, Xu S, et al. Silencing CDCA5 induces apoptosis. J Clin Lab Anal. 2022;36(5):e24396.
32. Lee Y, Wu Z, Yang S, Schreiner SM, Gonzalez-Smith LD, Rhie SK. MYBL2-regulated genes in lung adenocarcinoma. Cancers. 2022;14(20):4979.
33. Holland SJ, Pan A, Franci C, Hu Y, Chang B, Li W, et al. R428 blocks Axl and tumor spread. Cancer Res. 2010;70(4):1544–55.
34. Lee KW, Park YS, Ahn JB, Rha SY, Kim HK, Lee PY, et al. Vactosertib plus pembrolizumab in metastatic cancer. J Immunother Cancer. 2019;7(Suppl 1):P377.