Heliyon 10 (2024) e29126

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Identification of CNKSR1 as a biomarker for "cold" tumor microenvironment in lung adenocarcinoma: An integrative analysis based on a novel workflow

Qidong Cai [a,b], Mou Peng [a,b,c,*]

[a] Department of Thoracic Surgery, The Second Xiangya Hospital, Central South University, Changsha, Hunan, 410011, China
[b] Hunan Key Laboratory of Early Diagnosis and Precision Therapy, Department of Thoracic Surgery, The Second Xiangya Hospital, Central South University, Changsha, Hunan, 410011, China
[c] Department of Urology, The Second Xiangya Hospital, Central South University, Changsha, Hunan, 410011, China

ARTICLE INFO

Keywords:
CNKSR1
PD-L1
Lung adenocarcinoma
Immunotherapy
Machine learning
Tumor microenvironment

ABSTRACT

Background: Therapies targeting PD1/PD-L1 pathway have revolutionized the treatment of lung cancer. However, anti-PD1/PD-L1 therapies have proven beneficial for only a select group of lung adenocarcinoma (LUAD) patients and generally do not work for immuno-cold tumors characterized by a lack of immune cell infiltration. Identifying novel biomarkers is vital to broad therapeutic options for LUAD patients with no response to anti-PD1/PD-L1 immunotherapies.
Methods: Our study has developed a novel strategy to identify a promising biomarker that addresses the limitations of anti-PD1/PD-L1 immunotherapy in treating immunological cold tumors. We exacted LUAD RNA-seq data from the Cancer Genome Atlas database (TCGA). Using several machine learning methods, we identified the candidate biomarker. Based on the expression level of PD-L1 and the identified biomarker, samples were categorized into four groups. We further used ESTIMATE, ssGSEA, and CIBERSORT algorithms to calculate the immune infiltration level of each group. The results were validated in three independent bulk datasets and one scRNA-seq dataset. Immunohistochemistry (IHC) assessments were performed in clinical samples to further evaluate the coexpression of CNKSR1 and PD-L1, and to compare CD8 + T cell infiltration among groups.
Results: After comprehensive analyses, CNKSR1 was identified as a novel promising biomarker for immuno-cold LUAD. CNKSR1 mRNA expression levels exhibited a negative correlation with both PD-L1 mRNA expression and the extent of immune cell infiltration in LUAD. Besides, in contrast to the significant association between the expression of PD-L1 and the majority of other well-established or widely studied immune checkpoint molecules, a mutually exclusive expression pattern is observed between CNKSR1 and these molecules. The aforementioned results were consistent in validation datasets. The prognostic model built based on the CNKSR1 coexpression module also showed robust predictive performance. Additionally, IHC assessments have confirmed that the coexpression of CNKSR1 and PD-L1 is rare in LUAD samples. Notably, LUADs in the high-CNKSR1 group, characterized by high CNKSR1 but low PD- L1 expression, demonstrated reduced infiltration of $CD8^+$ T cells.

* Corresponding author. Department of Thoracic Surgery, The Second Xiangya Hospital, Central South University, Changsha, Hunan, 410011, China.
E-mail address: pengmou@csu.edu.cn (M. Peng).

https://doi.org/10.1016/j.heliyon.2024.e29126
Received 2 February 2024; Received in revised form 20 March 2024; Accepted 1 April 2024
Available online 4 April 2024

Conclusions: In summary, CNKSR1 emerges as a promising biomarker for immune-cold LUADs, and the study into CNKSR1 modulating T-cell infiltration may lead to the identification of compensatory molecules to enhance the effectiveness of current immunotherapy for LUAD.

1. Introduction

Immunotherapy has ushered in significant advancements in the treatment of advanced or metastatic non-small cell lung cancer (NSCLC) in recent years [1]. PD1/PD-L1 immune checkpoint inhibitors (ICIs) have already shown promising therapeutic outcomes and become the first-line treatment for advanced NSCLC patients [2,3]. However, it is important to note that only a relatively small percentage of advanced NSCLC patients have experienced tangible benefits from ICI therapy. PD1/PD-L1 monotherapy, for instance, has consistently yielded an objective response rate (ORR) of approximately 20% in unselected NSCLC patients [4–8]. Among patients with a PD-L1 tumor proportion score (TPS) of 50% or higher, there has been a notable improvement in overall survival, with 29.6% of treatment-naive and 25.0% of previously treated NSCLC patients deriving substantial benefits in overall survival [9]. This observation underscores the fact that the PD1/PD-L1 axis represents only a fraction of the broader immune response against cancer. Consequently, multiple checkpoint molecules, such as IDO1, LAG3, TIM3, SIGLEC15, and others, have been identified and demonstrated encouraging prospects in clinical application [10,11]. Inhibiting these checkpoint proteins has been shown to provide complementary advantages when used alongside PD1/PD-L1 blockade therapy or exert antitumor effects in PD-L1 negative cases of human cancers.

The effectiveness of ICIs in combating tumors hinges on the levels of PD-L1 expressed within the tumors and the ability of effector immune cells to recognize and eliminate cancer cells. The remarkable success of ICIs in NSCLC has spurred a surge in research, underscoring the growing necessity to characterize the immune landscape within the tumor-associated microenvironment. Tumor-infiltrating immune cells constitute a vital component and provide a reflection of the actual state of the tumor immune microenvironment [12]. Generally, tumors exhibiting high levels of effector immune cell infiltration and increased PD-L1 expression are classified as “hot” tumors [13]. In clinical practice, this phenotype tends to exhibit greater responsiveness to ICIs [14]. In contrast, the immunotherapy of “cold” tumors is an enduring conundrum due to the exclusion of immune cells. Numerous efforts have been dedicated to unraveling the underlying mechanisms and exploring therapeutic strategies for these immune “cold” tumors [13]. A prior study demonstrated that B7–H4 expression was mutually exclusive with PD-L1 in gliomas, and elevated B7–H4 expression could identify “cold” tumors. The discovery suggests the promising role of B7–H4 in glioma immunotherapy, offering a complementary option alongside PD- L1 [15]. Motivated by this insight, we embarked on a quest to identify the potential new biomarker for PD-L1 negative NSCLC patients.

In this study, we devised a comprehensive workflow to identify a biomarker exhibiting a negative correlation with both PD-L1 and immune cell infiltration in lung adenocarcinoma (LUAD). Leveraging genome-wide analyses based on data from The Cancer Genome Atlas (TCGA) database, we identified CNKSR1 as a gene with high expression levels that could effectively distinguish the immuno-cold subtype within LUAD. To enhance the credibility of our findings, we undertook a validation process involving three independent bulk RNA-seq datasets, one scRNA-seq dataset, and conducted immunohistochemistry (IHC) assessments using tissue samples. Our collective results solidify the significance of CNKSR1 as a pivotal component within the immune escape mechanism in LUAD.

2. Materials and methods

2.1. Data collection and preprocessing

We extracted LUAD bulk mRNA expression data with corresponding clinical information from TCGA database, Gene Expression Omnibus (GEO) databases (GSE50081, GSE40419), and IMvigor 210 cohort [16] (Atezolizumab treatment of bladder cancer research). Additionally, single nucleotide polymorphism (SNP) data from TCGA database were downloaded to acquire the mutation profiles of samples. RNA-seq data from TCGA and GSE40419 were downloaded in fragments per kilobase per million (FPKM) or reads per kilobase per million (RPKM) formats. mRNA data in GSE50081 were sequenced using a microarray-based sequencing approach. The TCGA-LUAD dataset encompassed a total of 526 samples of LUAD and 59 associated normal tissues, GSE40419 included 87 LUAD and 77 normal tissues, GSE50081 included 126 LUAD tissues, and IMvigor 210 cohort included 284 bladder cancer samples with immunological phenotype information verified by histopathological assessment. For the purposes of subsequent analyses, all the aforementioned RNA-seq data were converted to transcripts per million (TPM) values. In addition, scRNA-seq data from Bischoff et al. [17] including ten normal LUAD and ten normal lung tissues were obtained.

2.2. Bulk transcriptome data analysis

2.2.1. Marker identification

TCGA-LUAD transcriptome data were employed to identify a biomarker that met the following criteria: (1) Demonstrated high expression in LUAD tissues, (2) Showed mutual exclusivity with PD-L1 expression, and (3) Displayed increased expression associated with poor lymphocyte infiltration. The differentially expressed gene (DEG) analysis was performed through the Limma package. DEGs were identified based on an false discovery rate (FDR) < 0.05 and a log fold change >1. Subsequently, Spearman correlation analysis was applied to assess the correlations between PD-L1 and genes in TCGA-LUAD dataset. The intersection of the up-regulated DEGs was

taken with the genes which exhibited a correlation coefficient $r < -0.1$ and an FDR <0.05 for further analyses. The R package ESTIMATE (Estimation of STromal and Immune cells in MAlignant Tumour tissues using Expression data) was utilized to calculate stromal scores, immune scores, and ESTIMATE scores of each sample [18].

Feature selection was performed using WEKA software to determine the most crucial genes associated with ESTIMATE scores. Multiple methods, including information gain, gain ratio, symmetrical uncertainty, and reliefF were utilized. During this process, ESTIMATE scores were converted into categorical variables, and categorized into three subgroups based on tertile points. Each algorithm was executed with a 10-fold cross-validation approach. Finally, we took the intersection of the top 15 scoring genes obtained by the four algorithms as candidate genes.

2.2.2. Immune infiltration analysis

The ESTIMATE algorithm was applied to compute stroma, immune, ESTIMATE scores, as well as tumor purity for the samples [18]. For LUAD samples with available genomic sequencing data from TCGA, additional tumor purity data calculated using the ABSOLUTE algorithm based on somatic DNA alterations were also obtained from prior research [19,20]. Single sample gene set enrichment analysis (ssGSEA) was used to assess the infiltrating scores for 28 tumor-infiltrating lymphocytes (TILs). To mitigate potential bias arising from a single algorithm, CIBERSORT (Cell-type Identification By Estimating Relative Subsets Of RNA Transcripts) algorithm was utilized to calculate the absolute cell fractions for 22 different immune cell populations [21]. The results were subsequently corrected by tumor purity of each sample to obtain the final cell proportions.

2.2.3. Analysis based on clinical traits

Log-rank analysis was conducted on the TCGA dataset to investigate the prognostic significance of CNKSR1 in lung adenocarcinoma. Additionally, the relationships between CNKSR1 expression and clinical parameters, including AJCC stage, T stage, N stage, M stage, EGFR mutation status, and KRAS mutation status, were analyzed.

2.2.4. Gene coexpression module identification, functional annotation, and prognostic model construction

To learn about the underlying biological roles of CNKSR1, we identified CNKSR1-related gene coexpression module using biweight midcorrelation (bicor) [17]. Genes with the top 1% bicor values were included in the module. Then, Gene Ontology (GO) enrichment analysis assessed the biological processes (BP), molecular functions (MF), and cellular components (CC) of the module. Subsequently, to further evaluate the prognostic impact of the CNKSR1-related module, we utilized univariate Cox regression analysis and least absolute shrinkage and selection operator (LASSO)-penalized Cox regression to identify key prognostic genes and built a risk score formula to predict the prognosis of LUAD patients. Log rank test was used to determine the survival difference between the low- and high-risk groups, and the receiver operating characteristic (ROC) curve was used to evaluate the model performance. The prognostic model was trained in TCGA dataset, and validated in GSE50081 dataset. The two datasets were normalized to eliminate batch effects using sva package in this procedure.

2.3. Single-cell RNA sequencing data analysis

If not stated otherwise, scRNA-seq data were analyzed using Seurat v5 [22]. The data preprocessing, filtering, and normalization schemes were followed previous study by Bischoff et al. [17]. In more detail, raw scRNA-seq gene expression data of all patients were merged and converted to a Seurat object. Cells were retained if they had 500–10,000 genes detected, 1000–100,000 unique molecular identifiers (UMIs) counted, less than 30% mitochondrial reads, and less than 5% hemoglobin reads. Then the expression matrix was normalized with SCTransform function using the default parameters. We performed principal component analysis (PCA), and the top 20 principal components were used as input to constructed the shared nearest neighbor (SNN) graph. The SNN graph was further embedded in two-dimensional space using Uniform Manifold Approximation and Projection (UMAP). Main cell clusters were annotated using canonical cell type markers. For epithelial cells, we performed CopyKAT [23] analysis to infer the genomic copy number. Clusters with aneuploidy alterations in more than half of the cells were annotated as tumor cells. Epithelial cells were classified into subpopulations based on marker genes (Supplementary Table 1). Immune cell clusters were firstly identified by cell type markers of Habermann et al. [24]. Subsequently, the T cell cluster was subdivided into more detailed subgroups using CellTypist [25] algorithm in Python 3.7.

2.4. Clinical samples

The study adhered to the ethical guidelines of the 1975 Declaration of Helsinki and was approved by the Ethics Committees of the Second Xiangya Hospital of Central South University (2020-609). During September 2020 and January 2022, a total of 51 paraffin-embedded LUAD specimens that had been surgically removed at our department were obtained. Prior to participating in the study, all patients provided their written informed consent to participate and publish by signing the necessary documents. Detailed clinicopathological characteristics are provided in Supplementary Table 2.

2.5. Protein immunohistochemistry (IHC) and semi-quantitative analysis

The processing of paraffin sections obtained from LUAD patient samples involved several steps. Initially, these sections were deparaffinized and hydrated using xylene and ethanol. Subsequently, the deparaffinization and hydration of these sections were

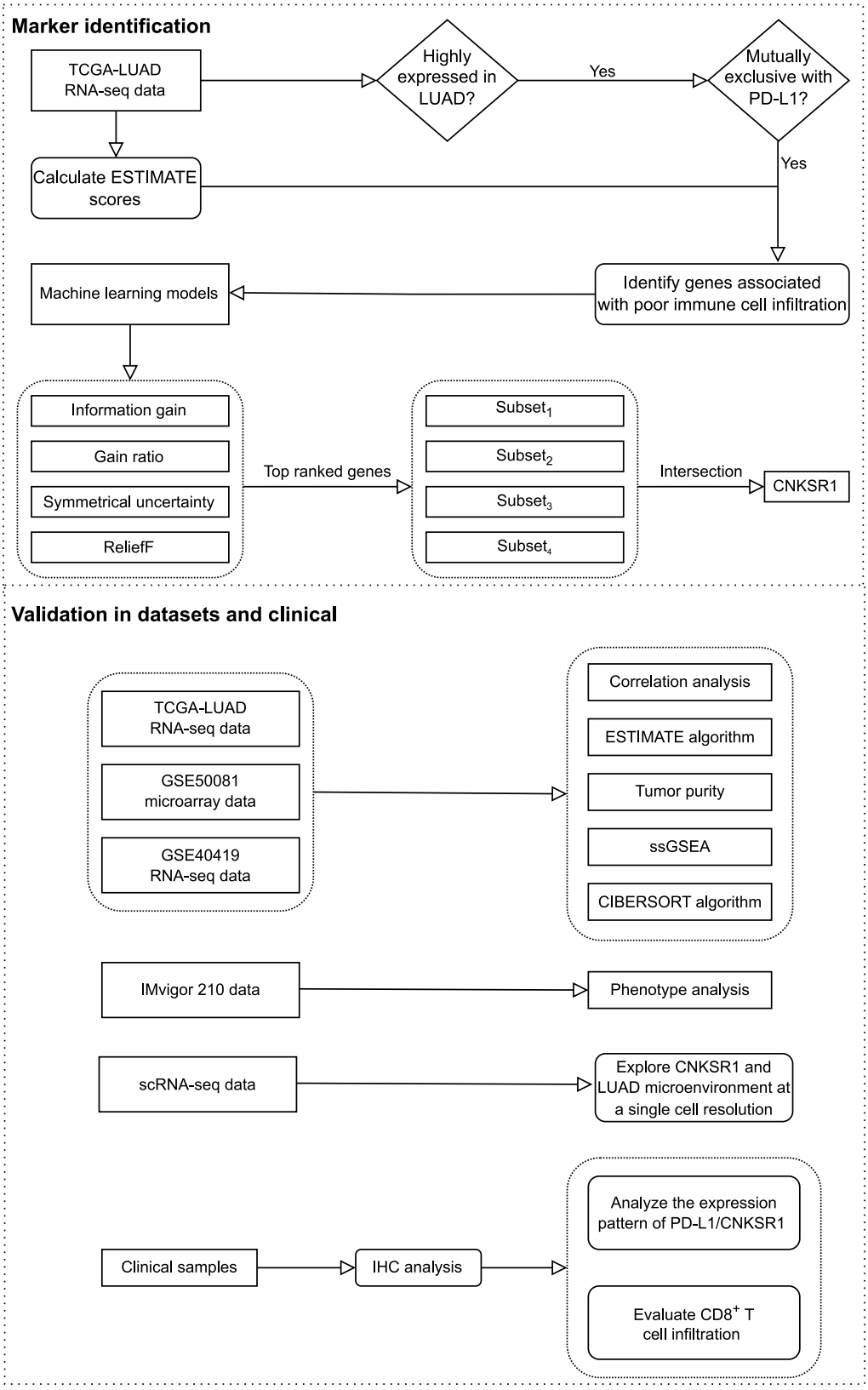

Fig. 1. The flowchart of the current study.

carried out using xylene and ethanol. After that, the antigen retrieval process was carried out using a solution of Tris-Ethylene Diamine Tetraacetie Acid (EDTA, pH = 9.0). The sections were subjected to incubation with specific antibodies, including anti-CNKSR1 antibody (#10885-1-AP; ProteinTech, Wuhan, China) at a 1/1000 dilution, anti-PD-L1 antibody (#66248-1-Ig; ProteinTech, Wuhan, China) at a January 2000 dilution, and anti-CD8 antibody (#RMA-0514; MXB Biotechnology, Fuzhou, China) using a working solution. Then the slides were washed extensively three times using PBS solution, with each wash lasting for 2 min, followed by a 15-min incubation with the corresponding *MaxVision* kit (#KIT-5001 or #KIT-5004; Maixin Biol, Fuzhou, China) at room temperature. Afterward, the slides were thoroughly rinsed with PBS and then dehydrated using gradient ethanol concentrations. Finally, we controlled the whole slides using xylene and coverslipped. Microscopic imaging of the slides was carried out using a CX41 microscope (OLYMPUS, Tokyo, Japan), equipped with the Microscope Digital Camera System DP-72 (OLYMPUS, Tokyo, Japan). The obtained images were surveyed and captured for subsequent analysis. IHC staining for CNKSR1 and PD-L1 was assessed semi-quantitatively using the immunoreactivity score (IRS) criterion, which considered both the staining intensity and the percentage of staining [26]. Based on the intensity of staining, CNKSR1/PD-L1 expression levels were manually graded into four groups: negative (0), weak (1), moderate (2), and strong (3). The percentage scores reflected the coverage percentage of immunoreactive tumor cells and ranged from 0 to 4, corresponding to 0%, 1–25%, 26–50%, 51–75%, and 76–100% coverage, respectively. Overall IHC scores were calculated by multiplying the intensity and percentage scores and ranged from 0 to 12. Low and high expression of CNKSR1/PD-L1 were defined as IHC score ≤ 4 and > 4, respectively. For assessing the extent of CD8 positive ($CD8^+$) cells, a semiquantitative scoring method, as previously published was used [27]: 1 for absence or scarce $CD8^+$ cells (less than five $CD8^+$ cells throughout the entire section); 2 for $CD8^+$ cells accumulated in localized field of view (Only single areas had $CD8^+$ cells infiltration); 3 for infiltration of $CD8^+$ cells throughout the entire section, but scattered sporadically; and 4 for high quantities of $CD8^+$ cells accumulated throughout the entire section. The immunohistochemical evaluation and scoring were performed by two independent pathologists.

2.6. Statistical analysis

All statistical analyses were conducted using R version 4.1.1. Default parameters were used for the algorithms unless specifically mentioned otherwise. Correlation analysis was performed using Spearman's rank correlation coefficient. For categorical variables, the $\chi 2$ test was applied, while differences in continuous variables between groups were assessed using either the Wilcoxon rank-sum test or Kruskal-Wallis test. Repeated measures were corrected using the Benjamini and Hochberg FDR method. Statistical significance was established with a two-tailed P-value <0.05 unless specified otherwise.

3. Results

3.1. Identification of a biomarker associated with cold tumors in LUAD

The flowchart of this study was illustrated in Fig. 1. Initially, a total of 5969 up-regulated and 4194 down-regulated DEGs were

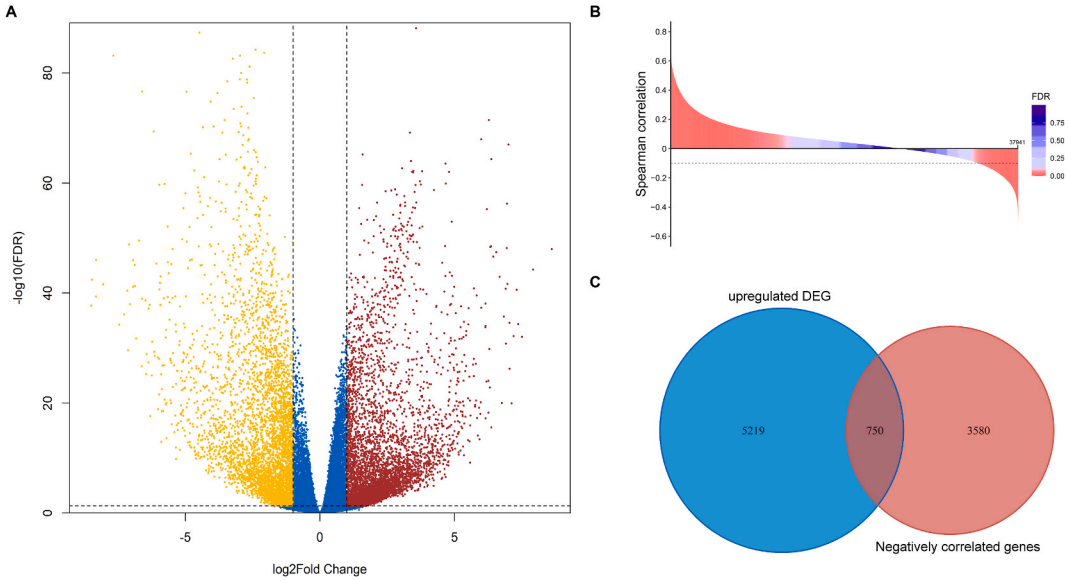

Fig. 2. CNKSR1 was identified as the candidate biomarker. (A) Volcano plot of DEGs. (B) Overview of correlation coefficients (y-axis) and FDR values (color depth) of Spearman correlation analysis. A total of 37941 genes were included in the diagram. (C) Venn diagram of the upregulated DEGs and genes significantly negatively correlated with PD-L1. FDR, false discovery rate. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

identified in TCGA dataset (Fig. 2A, Supplementary Table 3). Spearman correlation analysis identified 4330 genes that exhibited a noteworthy negative correlation with PD-L1 (Fig. 2B–Supplementary Table 4). Subsequently, 1144 DEGs were intersected for further analysis (Fig. 2C). To pinpoint the most promising biomarker, four distinct feature selection approaches were employed: information gain (IG), gain ratio (GR), symmetrical uncertainty (SU), and reliefF (RF). These approaches were designed to select a biomarker that not only displayed a negative correlation with PD-L1 but also served as an indicator of limited immune cell infiltration within the LUAD immune microenvironment. Subsequently, the top 15 ranked genes identified by each approach were considered, and their intersections were used to identify the final biomarker, as presented in Table 1. Ultimately, through this comprehensive approach, CNKSR1 was selected as the candidate biomarker with the most promising potential for association with cold tumors in LUAD.

3.2. High-CNKSR1 identifies immuno-cold tumors

In accordance with the classification method used in a prior study [28], high-expressed samples were defined as those with the gene exceeding one standard deviation above the mean expression level. For datasets with tumor sample numbers smaller than 100, to mitigate the impact of limited size of samples with high expression on subsequent statistical analyses, we classify high expression samples as those ranking in the top 25% of gene expression levels. Consequently, the samples within each dataset were categorized into four groups: high-CNKSR1, high-PD-L1, double-low, and double-high, based on the expression levels of CNKSR1 and PD-L1. Among the 526 patients in the TCGA dataset, 400 (76.05%) were assigned to the double-low group, 80 (15.21%) were in the high-CNKSR1 group, 43 (8.17%) were categorized as the high-PD-L1 group, and only 3 (0.57%) were in the double-high group Supplementary Fig. 1A. The coexpression of PD-L1 and CNKSR1 was found to be limited. In addition to the mutually exclusive relationship between PD-L1 and CNKSR1, our analysis revealed that most other well-established or widely studied checkpoint immune molecules, exhibited a significant association with PD-L1 while displaying mutually exclusive expression with CNKSR1 (Fig. 3A). Besides, the expressions of most of these checkpoint molecules were lowest in high-CNKSR1 group, while highest in the high-PD-L1 group (Fig. 4). Furthermore, when assessing the results analyzed by ESTIMATE, it became evident that LUAD samples in the high-PD-L1 group exhibited the highest level of immune infiltration, whereas the high-CNKSR1 group displayed the lowest level of immune infiltrating scores (Fig. 3B). The high-CNKSR1 group demonstrated the highest tumor purity, as supported by transcriptome data (Fig. 3C) and somatic mutation data (Supplementary Fig. 1B). This higher tumor purity contributed to the reduced tumor immune activity observed in this group. Furthermore, the ssGSEA score, which indicates the extent of infiltration by 28 immune cell types, suggested the least TILs infiltration in the high-CNKSR1 group (Fig. 3D). CIBERSORT evaluation reflecting the absolute proportions of different types of immune cells corroborated that tumors in the high-PD-L1 group had the highest proportion of $CD4^+$ T cells, $CD8^+$ T cells, and tumor-associated macrophages (TAMs) infiltration, whereas the high-CNKSR1 group had the least (Fig. 5A).

The above results were further validated in two independent datasets: GSE50081 and GSE40419. Notably, the GSE40419 dataset included RNA-seq data of 77 adjacent normal tissues in addition to LUAD samples. CNKSR1 exhibited significantly higher expression levels in LUAD tissues compared to normal tissues (Supplementary Fig. 2B). The results consistently demonstrated that CNKSR1 mRNA expression was significantly mutually exclusive with PD-L1 and most other checkpoint immune molecules (Supplementary Figs. 3A and 4A). Patients displaying co-high expression of CNKSR1 and PD-L1 were rare in both datasets. Among the 127 patients in the GSE50081 dataset, 85 (66.93%) fell into the double-low group, 20 (15.75%) were classified as high-CNKSR1, 21 (16.54%) were high-PD-L1, and only 1 (0.79%) was in the double-high group (Supplementary Fig. 2A). Among the 87 patients in the GSE40419 dataset, 46 (52.87%) belong to the double-low group, 19 (21.84%) to high-CNKSR1, 19 (21.84%) to high-PD-L1, and 3 (3.45%) were in the double-high group (Supplementary Fig. 2C). Similarly, we conducted ESTIMATE analysis, ssGSEA analysis, and CIBERSORT analysis in the two validation datasets, which consistently indicated that the high-CNKSR1 group LUADs had a deficiency of TILs especially $CD8^+$ T cells in the tumor microenvironment (TME) when compared with the other two groups (Supplementary Figs. 3 and 4; Fig. 5B and C).

Although the profiles of immune cell infiltration described above were analyzed using multiple reliable algorithms, these results

Table 1
Top 15 ranked genes identified by different feature selection methods for ESTIMATE scores in TCGA-LUAD dataset.

SU	RF	IG	GR
OVOL2	RAB25	OVOL2	CPE
CNKSR1	DHODH	GSTP1	FKBP2
GSTP1	OVOL2	POLR1C	RAB25
POLR1C	TM7SF2	GRHL2	CLDN7
C6orf136	GSTP1	MRPS26	NAXE
CLDN7	PRSS8	RCCD1	LINC02313
C2orf15	ELMO3	C2orf15	ICA1
GRHL2	PPP1R1B	C6orf136	TXNL4A
MRPS26	MANBAL	MRPS9	HGD
BLOC1S4	TCEA3	**CNKSR1**	C6orf136
NELFA	TMC4	PITPNA-AS1	MRPS18A
RAB25	AKAP1	PET117	**CNKSR1**
MTA3	**CNKSR1**	EPCAM	SOX12
PITPNA-AS1	SLC22A31	BLOC1S4	GRTP1
PET117	C2orf15	KRTCAP3	MTA3

Abbreviation: SU, symmetrical uncertainty. RF, relief. IG, information gain. GR, gain ratio.

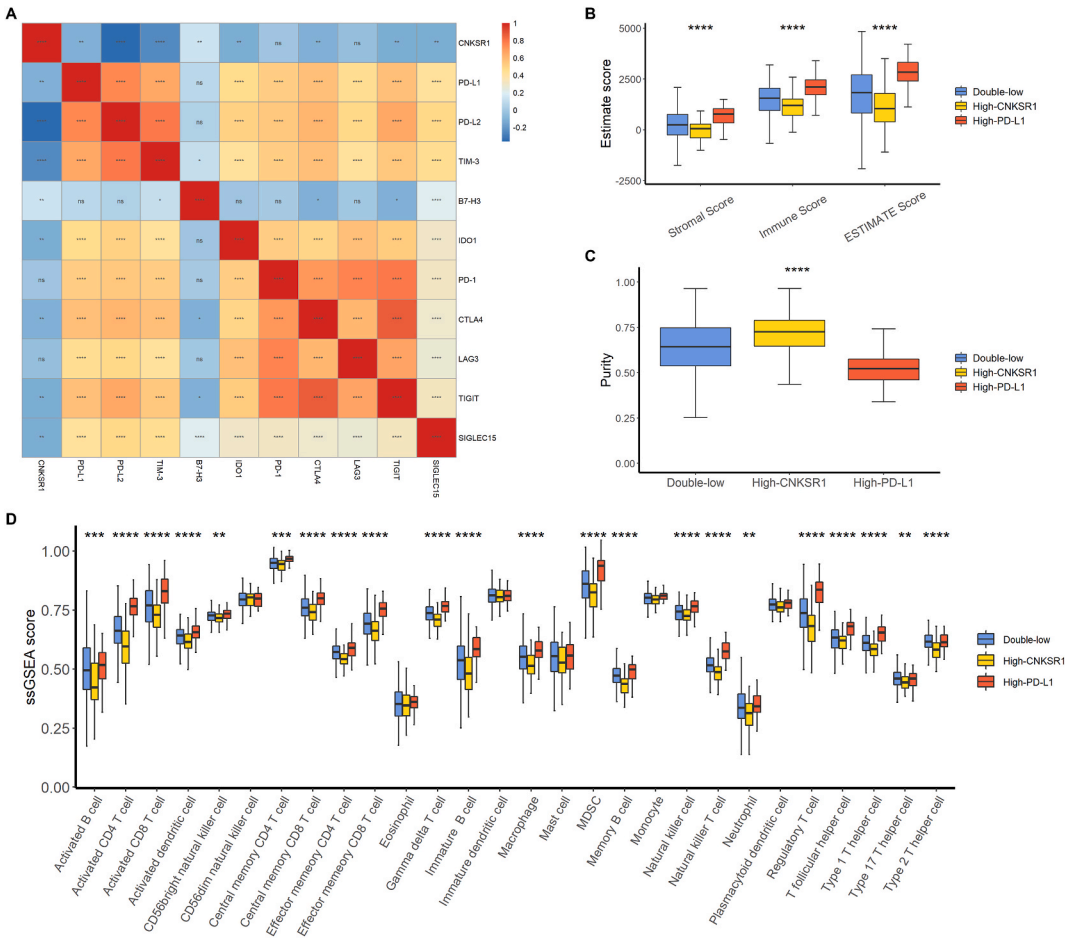

Fig. 3. The analysis of TCGA-LUAD dataset revealed the correlation between CNKSR1 and PD-L1 as well as other checkpoint immune molecules, and the immune infiltration profile of LUAD samples according to the PD-L1/CNKSR1 classifier. (A) Correlations among CNKSR1 and immune checkpoint factors. P values in the Spearman correlation matrix were FDR-corrected. (B) Stromal score, immune score, and ESTIMATE score among different groups evaluated by ESTIMATE algorithm. (C) Differences in tumor purity among groups. (D) ssGSEA evaluation demonstrated the infiltration differences of different types of immune cells among double-low, high-CNKSR1, and high-PD-L1 groups. In B-D, the comparisons were conducted among three groups by Kruskal-Wallis test. *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001.

were obtained through mathematical models rather than histopathological assessments. However, in the IMvigor 210 cohort of bladder cancer, which comprises 284 samples with histopathologically verified tumor-immune phenotype, we analyzed the distribution of immunological phenotypes among patients in the high-CNKSR1, high-PD-L1, and double-low groups. Significant differences in tumor-immune phenotype were observed among these groups (Supplementary Fig. 5). Consistent with the aforementioned analyses, the high-CNKSR1 group displayed the poorest $CD8^+$ T cell infiltration, with the highest proportion of immune desert phenotype (46.81%), while the high-PD-L1 group demonstrated the most abundant $CD8^+$ T cell infiltration with the highest proportion of immune inflamed phenotype (60.71%).

3.3. Results validation at single-cell resolution

We analyzed scRNA-seq data to explore the expression pattern of CNKSR1 in LUAD, and its relationship with tumor immune microenvironment. After data preprocessing and filtering, a total of 114489 single-cell transcriptomes were subjected to further analysis. Samples of different origins were clustered and visualized by UMAP (Fig. 6A). After cell population annotation using canonical marker gene, 19517 epithelial, 90735 immune, and 4237 stromal transcriptomes were covered (Fig. 6B). As shown in Fig. 6C and D, CNKSR1 was expressed predominantly in cells of epithelial origin, and was scarcely expressed in lung immune and stromal cells. We further reclustered epithelial cell population. Epithelial cells were divided into diploid and aneuploid clusters by CopyKAT algorithm (Supplementary Fig. 6A), followed which malignant cells were discriminated from normal cells (Fig. 6E). As for normal cells, we identified alveolar type 1 (AT1), alveolar type 2 (AT2), basal, club, cilated, goblet, and neuroendocrine cells (Fig. 6E). Fig. 6F and G depicted the expression of CNKSR1 and PD-L1 in normal epithelial and tumor cells. CNKSR1 showed comparable expression levels in

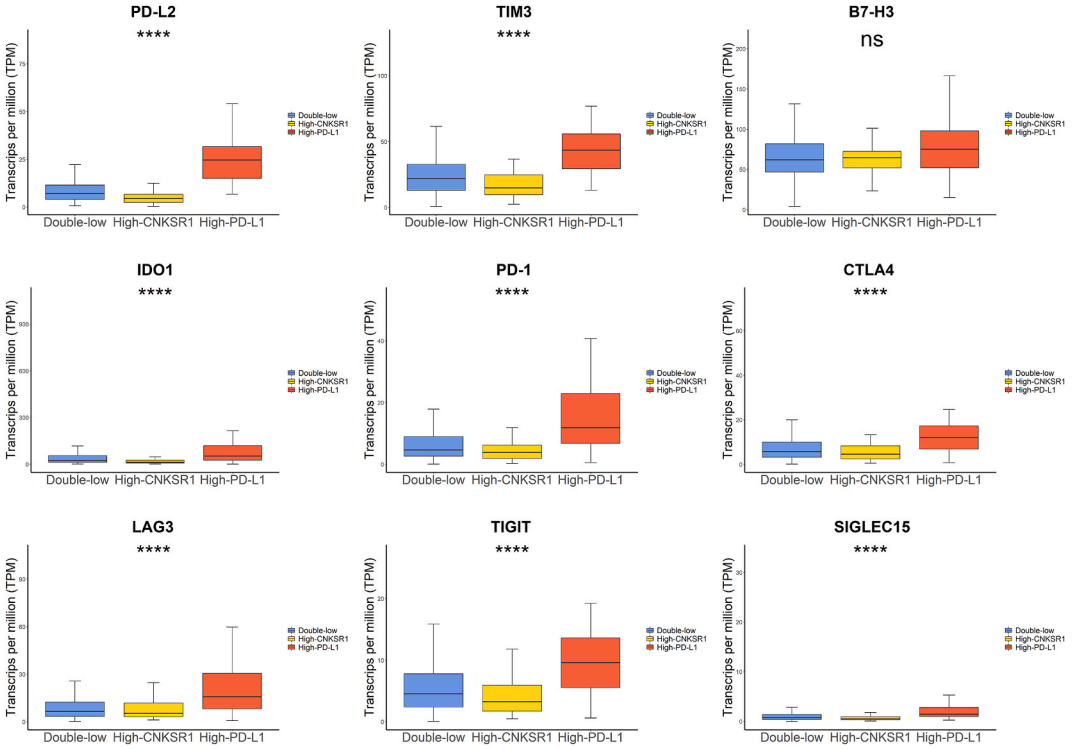

Fig. 4. Boxplot showing the differential expression of important checkpoint immune molecules in double-low, high-CNKSR1, and high-PD-L1 groups. The comparisons were conducted among three groups by Kruskal-Wallis test. ****P $<$ 0.0001; ns, not significant.

different types of epithelial cells (Fig. 6F). The coexpression pattern of CNKSR1 and PD-L1 in LUAD cells was displayed in Fig. 6H, indicating a mutually exclusive expression between CNKSR1 and PD-L1. After removing LUAD cell without expression of both genes, a strong negative correlation (R = −0.959, P $<$ 2.2e-16) was shown between the two genes (Supplementary Fig. 6B). Besides, compared with normal samples, the average expression levels of CNKSR1 were higher in LUAD cells (Fig. 6I).

To assess the tumor immune microenvironment, immune transcriptomes were reclustered and annotated (Fig. 6J). T cell population was further subdivided and $CD8^+$ T cells were identified using CellTypist algorithm. We performed pseudobulk analysis to calculate the aggregated expression of CNKSR1 in each LUAD sample, and the correlation analysis between CNKSR1 expressions and the fractions of $CD8^+$ T cells in immune cells was performed. As shown in Fig. 6K, the aggregated expressions of CNKSR1 had a negatively correlated trend (R = −0.534, P = 1.12e-01) with the fractions of T cells. Using canonical immune signatures including the cytotoxic signature (GZMA, GZMB, GZMK, GNLY, IFNG, PRF1, and NKG7) and the exhausted signature (LAG3, TIGIT, PCCD1, HAVCR2, CTLA4, LAYN, and ENTPD1), we also calculated the correlation between CNKSR1 levels with the cytotoxic or exhausted score of T cells. Results showed the aggregated expression of CNKSR1 had strong negative correlation (R = −0.903, P = 8.8e-04) with the cytotoxic score, and represented a negatively correlated trend (R = −0.600, P = 7.31e-02) with the exhausted score, suggesting that CNKSR1 had a role in negatively regulating immune infiltration (Fig. 6K). To assess whether CNKSR1 exhibits an association with TAMs at the single-cell resolution, we analyzed the correlation between CNKSR1 expressions and the fractions of TAMs (R = 0.435, P = 2.09e-01; Supplementary Fig. 6C). Additionally, using the M1 and M2 macrophage signatures summarized by Bagaev et al., we further examined the correlations between CNKSR1 and M1/M2 macrophage signature scores (R = −0.503, P = 1.43e-01; R = 0.261, P = 4.37e-01; Supplementary Fig. 6C). However, no clear association between CNKSR1 and TAMs was observed in the scRNA-seq dataset.

3.4. Coexpression of PD-L1 and CNKSR1 is limited in LUAD clinical samples

In the cohort subjected to IHC analysis (N = 51), the distribution of LUAD patients based on their expression of PD-L1 and CNKSR1 revealed the following patterns: 23.53% had low expression of both PD-L1 and CNKSR1 (double-low group), 49.02% displayed high expression of PD-L1 but low expression of CNKSR1 (high-PD-L1 group), and 21.57% exhibited high expression of CNKSR1 but low expression of PD- L1 (high-CNKSR1 group). Remarkably, only 5.88% of the patients demonstrated high expression of both PD-L1 and CNKSR1 (double-high group) (Fig. 7A). Representative samples of these subgroups are depicted in Fig. 7B.

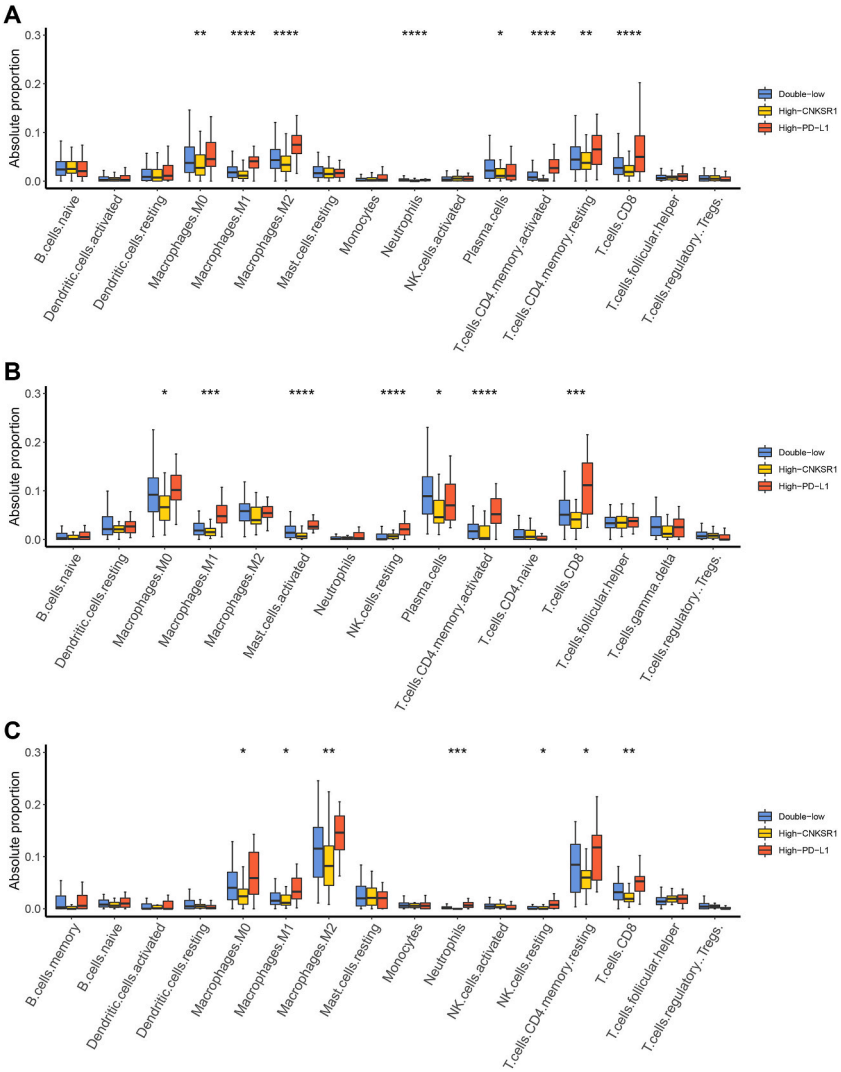

Fig. 5. Differences of absolute immune cell fractions for immune cell populations among high-CNKSR1, high-PD-L1, and double-low groups in (A) TCGA-LUAD, (B) GSE50081, and (C) GSE40419 datasets. Cells with no presence in more than half of the samples were not presented in the figures. The comparisons were conducted among three groups by Kruskal-Wallis test.

3.5. PD-L1/CNKSR1 classifier identified LUAD subtypes with different degrees of $CD8^+$ T cell infiltration

Given that $CD8^+$ T cells consistently exhibited higher infiltration in the high-PD-L1 group and lower infiltration in the high-CNKSR1 group in the previous analyses, and $CD8^+$ T cells are pivotal players in current cancer immunotherapy [29,30], we conducted IHC staining for $CD8^+$ T cells in LUAD clinical samples. Our findings revealed that LUAD samples in the high-CNKSR1 group exhibited the lowest levels of $CD8^+$ T-cell infiltration compared to both the high-PD-L1 group and the double-low group (Fig. 8). This discrepancy may contribute to the establishment of the TME of "cold" tumors.

3.6. Relationship of CNKSR1 and clinical traits

The Kaplan Meier survival curve analysis showed no significant correlation between CNKSR1 expression and survival time ($P = 4.38e\text{-}01$; Supplementary Fig. 7A). Additionally, there were no significant changes observed in CNKSR1 expression between subgroups divided by AJCC stage, T stage, N stage, M stage, EGFR mutation status, and KRAS mutation status ($P > 0.05$; Supplementary Fig. 7B). CNKSR1 expression was significantly higher in the EGFR mutation group ($P = 2.00e\text{-}02$; Supplementary Fig. 7B), conforming to the previous finding that CNKSR1 has close association with tyrosine phosphorylation [31,32]. This also suggests that CNKSR1 may be implicated in the immune-inert status of EGFR-mutant lung adenocarcinoma.

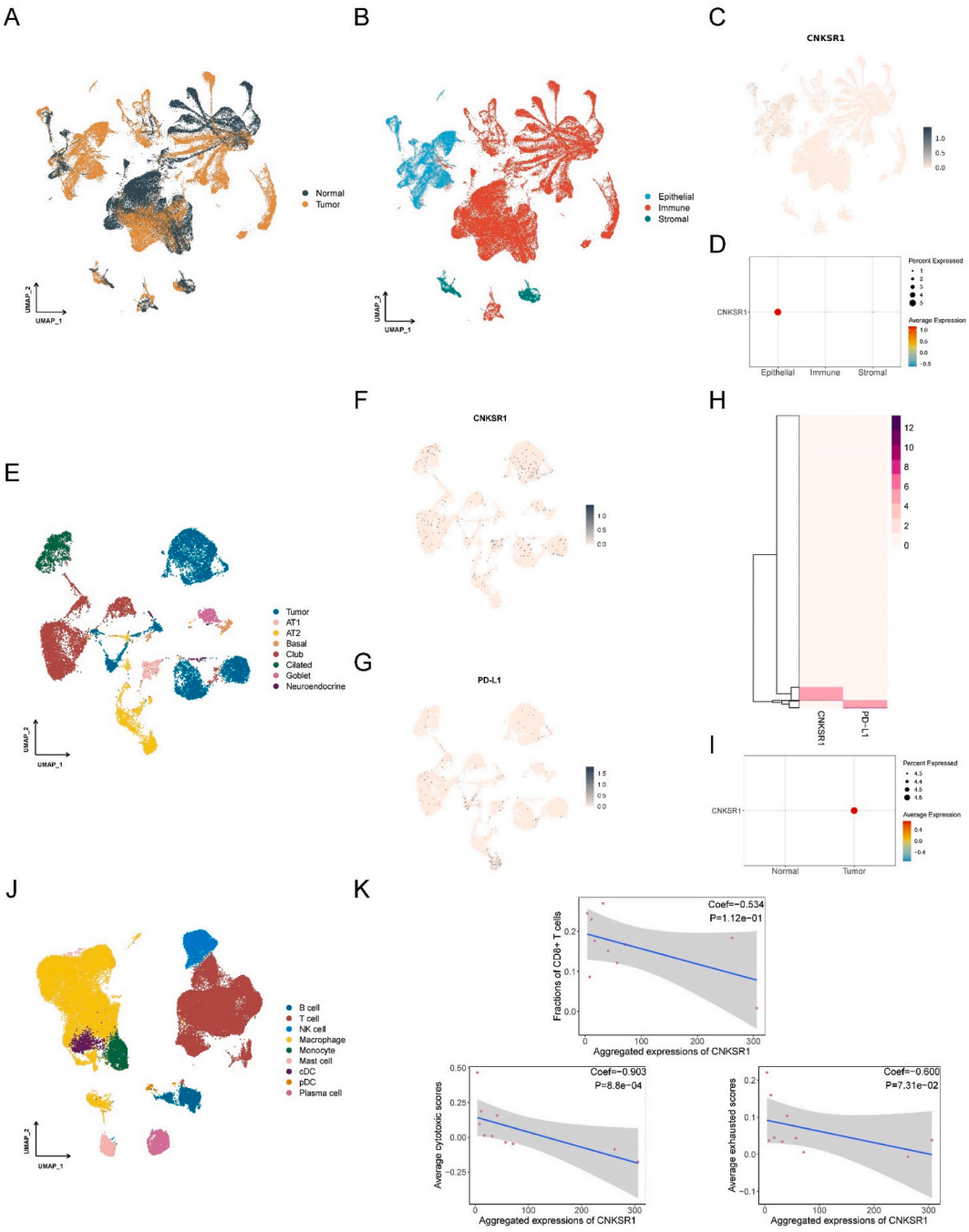

Fig. 6. scRNA-seq analysis revealed the expression pattern of CNKSR1 and its relationship with tumor immune microenvironment in high-resolution. (A, B) UMAPs of the 114489 cells, with each cell color-coded for: type of sample source (A), main cell type (B). (C) Expression and distribution of CNKSR1 in the 114489 cells visualized by FeaturePlot. (D) Expression levels of CNKSR1 in epithelial, immune, and stromal cells. (E) UAMP of the 19517 epithelial cells color-coded by cell type. (F, G) Expression and distribution of CNKSR1(F) and PD-L1(G) in all epithelial cells. (H) Heatmap represented the coexpression pattern of CNKSR1 and PD-L1 in tumor cells. (I) Expression levels of CNKSR1 in tumor and normal lung epithelial cells. (J) UAMP of the 90735 immune cells color-coded by cell type. (K) Correlations of the aggregated expressions of CNKSR1 with the fraction of $CD8^{+}$ T cells in immune cells (top), average cytotoxic scores (bottom left), and average exhausted scores (bottom right). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

A

PD-L1/CNKSR1 subgroups	PD-L1	CNKSR1	N (%)
Single high-expression (High-PD-L1)	High	Low	25 (49%)
Single high-expression (High-CNKSR1)	Low	High	11 (22%)
Double high-expression (Double-high)	High	High	3 (5%)
Double low-expression (Double-low)	Low	Low	12 (24%)

B

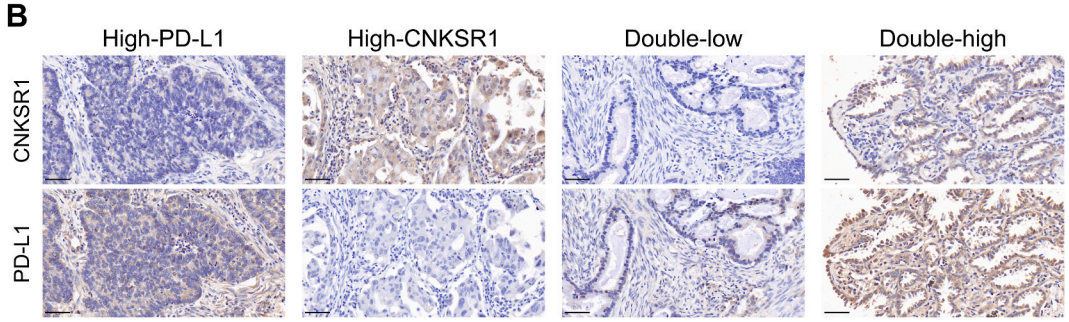

Fig. 7. IHC analysis showed both CNKSR1 and PD-L1 co-expressed at a high level was rare. (A) The distribution of the high-PD-L1, high-CNKSR1, double-high, and double-low groups in the IHC cohort. (B) Representative images of PD-L1/CNKSR1 groups. Magnification, × 400, Bar = 50 μm. IHC, Protein immunohistochemistry.

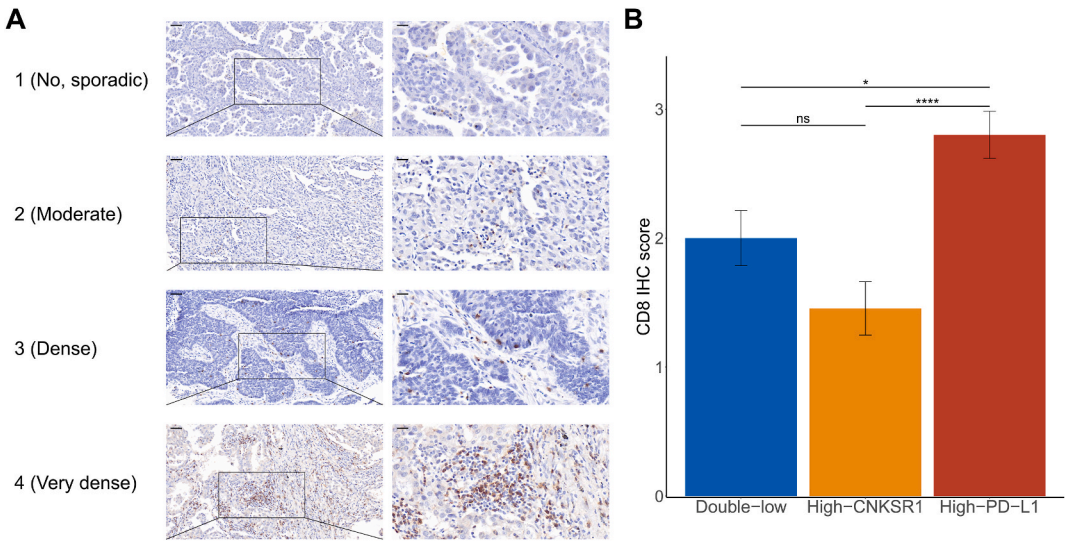

Fig. 8. Disparities in the infiltration of $CD8^+$ T cells among the high-CNKSR1, high-PDL1, and double-low groups. (A) Representative cases of $CD8^+$ T cell infiltration with IHC score. (B) The high-CNKSR1 group showed significantly less $CD8^+$ T cell infiltration than the high-PD-L1 group. Magnification, × 200, Bar = 50 μm *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001.

3.7. Identification of CNKSR1 coexpression module, functional enrichment analysis, and generation and validation of prognostic signature

A total of 204 genes (Supplementary Table 5) were identified in the CNKSR1-related coexpression module, and were subsequently mapped to the GO database. In the BP ontology, protein catabolic process, regulation of GTPase activity, and protein deacylation were the most enriched terms. In the MF ontology, small GTPase binding, serine/threonine kinase activity, and Ras guanyl−nucleotide exchange factor activity were the most enriched terms. In the CC ontology, site of polarized growth, SCF ubiquitin ligase complex, and recycling endosome were the most enriched terms (Fig. 9A).

To explore the prognostic impact for LUAD of these genes, we performed univariate Cox proportional hazards regression analysis, and 13 genes were selected as survival-related genes (Supplementary Table 6). Afterward, LASSO regression model was used to variable reduction. Finally, a ten-gene multivariate survival model was established (Supplementary Table 6). Using the optimal

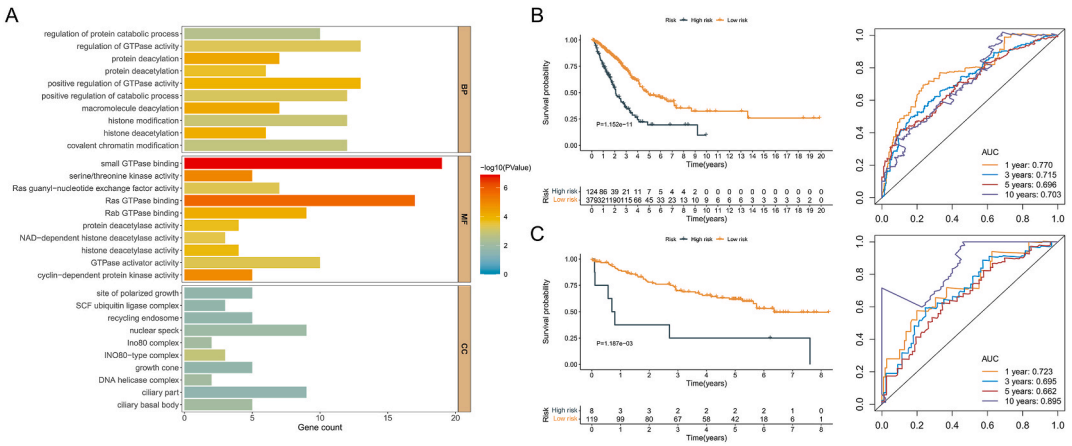

Fig. 9. Functional enrichment analysis and survival analysis. (A) Gene ontology (GO) analysis for CNKSR1 coexpression module based on biological process (BP), molecular function (MF), and cellular compartment (CC). (B) Kaplan-Meier curves (left) and ROC curves (right) of the prognostic model in the training set. (C) Kaplan-Meier curves (left) and ROC curves (right) of the prognostic model in the test set.

cutpoint identified by R package survminer, samples in the train set (TCGA-LUAD dataset) were divided into high- and low-risk groups. As shown in the Kaplan-Meier survival curves, patients in the high-risk group had higher mortality rates than those in the low-risk group ($P = 1.152e{-}11$; Fig. 9B). The prognostic performance of the model was measured by ROC curves, and the area under the curves (AUC) for 1, 3, 5, and 10 years were 0.770, 0.715, 0.696, and 0.703 (Fig. 9B), respectively. GSE50081 was used as the test set for the prognostic model, and the Kaplan-Meier survival curves showed LUAD patients with higher risk scores had lower survival probability than those with lower risk scores ($P = 1.187e{-}03$; Fig. 9C). The AUC of the test set for 1, 3, 5, and 10 years were 0.723, 0.695, 0.662, and 0.895, respectively, indicating the robustness of the signature. To precisely evaluate the predictive performance of the risk score model, we conducted a stratified log-rank analysis, considering factors such as age, smoking history, AJCC stage, lymph node metastasis, distant metastasis, KRAS mutation, and EGFR mutation. The results revealed that the high-risk group was consistently associated with unfavorable survival outcomes across all subgroups ($P < 0.05$; Supplementary Fig. 8).

4. Discussion

The absence of immune checkpoints and limited immune cell infiltration are critical factors that reduce the effectiveness of cancer immunotherapy [33]. In many cases, the expression of PD-L1 on LUAD cells serves as a crude measure of an ongoing tumor-specific immune response. Nonetheless, immune evasion mechanisms in solid tumors exhibit a significant degree of heterogeneity, with the PD1/PD-L1 pathway accounting for immune dysfunction in less than 40% of human solid tumors [34]. One potential solution for immunotherapy in PD-L1 negative cancers is to deepen our understanding of the mechanism governing tumor immune responses and explore new therapeutic targets. For instance, Chen et al. have identified Siglec-15 as an immune-suppressive molecule that is mutually exclusive to PD-L1 and is broadly upregulated on human cancer cells, making it a potential target for normalizing cancer immunotherapy [10]. Likewise, Zhou et al. discovered that PD- L1 and B7– H4 may function as mutually compensatory immune checkpoint molecules in gliomas, providing avenues for alternative immunotherapy strategies [15]. In recent years, the advancement of high-throughput sequencing has led to an explosion of transcriptomic information, enhancing our understanding of various cancers. Alongside this development, various algorithms for data mining have emerged. Notably, algorithms like ESTIMATE can accurately infer the fraction of stromal and immune cells in tumor samples based on gene expression data [18]. Increasingly more single-cell datasets being publicly available makes the results from bulk data well-validated. In our current study, we primarily employed three algorithms (ESTIMATE, ssGSEA, and CIBERSORT) to evaluate immune cell infiltration in the TCGA-LUAD dataset. We utilized several feature selection approaches to identify a promising biomarker that exhibited mutually exclusive expression with PD-L1 and had a close association with “cold” tumors. CNKSR1 emerged as a candidate biomarker, and subsequent validation in three datasets, one scRNA-seq dataset, and clinical samples highlighted its potential as a complement to PD-L1. The absence of a double-high subgroup (high expression of both PD- L1 and CNKSR1) in our study suggests the existence of unique immune evasion pathways within the LUAD subgroups we examined. Our findings offer a new perspective into the exploration of potential alternative immunotherapy targets in subgroups with low PD-L1 expression.

CNKSR1 is known as a multidomain scaffold protein [35,36]. Via its specific domains, CNKSR1 binds with various signaling molecules and is important for cell proliferation, survival, and migration [37]. CNKSR1 promotes the Ras/Raf/Mek/Erk and PI3K/AKT signaling cascade in a mutually exclusive manner upon different growth factor stimulation signal intensities [38,39]. At low signal intensity, CNKSR1 engages with CRAF, leading to the activation of ERK. In contrast, when the signal becomes more intense, CNKSR1 acts as a scaffold for AKT/RAF crosstalk, ultimately resulting in the AKT-driven phosphorylation of the inhibitory site within Raf [40]. Additionally, the acetylation of CNKSR1 serves as a positive feedback regulator for the Erk signaling pathway [40]. Furthermore, CNKSR1 collaborates with AKT to activate the alternative NF-κB pathway, leading to the upregulation of matrix metalloproteinase 14

(MMP14) expression, thereby enhancing cancer invasiveness [41]. Additionally, CNKSR1 mediates the initiation of the c-Jun N-terminal kinase (JNK) pathway by binding to Rho guanine nucleotide exchange factors, the GTPase Rho, and downstream molecules of Rho [26,42]. CNKSR1 has been implicated as an oncogene in various cancers [41,43,44], and its potential role in immunotherapy for bladder and liver cancers has also been suggested [45,46]. However, there has been a lack of in-depth investigation into the role of CNKSR1 in tumor immunity. Our study has revealed that the CNKSR1/PD-L1 classifier can effectively stratify "hot" and "cold" tumors in LUAD. We observed a decrease in $CD8^+$ T-cell infiltration in the high-CNKSR1 group, indicating that CNKSR1-mediated immune escape mechanisms may hold promise for the treatment of "cold" tumors. While extensive research has been conducted to identify potential immunotherapy biomarkers in cancers, many of the discoveries involve gene signatures composed of multiple molecules, and the expression levels of the signatures were parallel to that of PD-L1 [47,48]. In contrast, we believe that the identification of a specific immune-related gene that is mutually exclusive with PD-L1 holds greater significance for future research and application. The methodology employed in our current study offers a novel approach to uncover new immunological targets in cancers.

Various algorithms and feature selection methods were utilized in the current study, and it is necessary to provide more details about them. ssGSEA, ESTIMATE, ABSOLUTE, and CIBERSORT are all well-validated algorithms to infer cell fractions in tumor bulk data. ssGSEA is an extened algorithm of GSEA, capable of computing a resultant score based on the input gene expression data of a specific gene signature. The score reflects the overall upregulation or downregulation of the gene signature. Therefore, using cell-specific gene signatures, ssGSEA can estimate the abundance of individual cell types in bulk data [49]. Through the utilization of pre-defined signatures, ESTIMATE employs ssGSEA to calculate tumor purity and the proportions of stromal and immune cells from gene expression data [18]. Unlike ESTIMATE, ABSOLUTE algorithm utilizes copy-number alteration data to predict tumor purity, ploidy, and absolute copy numbers [19]. CIBERSORT is also a robust algorithm to infer the relative proportions of each cell type from mixtures of cells. It uses support vector regression (SVR) to accomplish the deconvolution of bulk tissue gene expression profiles based on reference expression signatures [19]. Integrating ssGSEA, ESTIMATE, ABSOLUTE, and CIBERSORT can enhance the robustness of the results from various aspects such as data sources and computational approaches. CopyKAT is an advanced Bayesian segmentation approach to estimate genomic copy number profiles from scRNA-seq data. An important application of CopyKAT is to discern malignant cells amidst a heterogeneous mixture of normal cells in scRNA-seq analysis. It achieves a high average genomic resolution of 5 Mb, facilitating the precise detection of copy number alterations with fine granularity [23]. CellTypist offers automated annotation of immune cell types and subtypes in scRNA-seq data. The machine learning model utilized by CellTypist was trained on cells originating from 20 distinct tissues sourced from 19 reference datasets [25]. This comprehensive training enhances the accuracy and robustness of cell type annotation across diverse tissues and organisms. Information gain, gain ratio, symmetrical uncertainty, and reliefF are major rank-based feature selection methods in WEKA software. Information gain, gain ratio, and symmetrical uncertainty ranked features using the concept of entropy drawn from information theory [50]. Meanwhile, reliefF assigned a score to each feature using the K-nearest neighbor method [51].

The limitations of our study should also be mentioned. Firstly, we did not explore the effects of CNKSR1 knockdown or overexpression on TILs in mouse models, as human CNKSR1 is not orthologous to mouse Cnksr1. Cause-effect experiments are crucial for validating the function of CNKSR1 in tumor immune evasion and assessing its potential as a target in immunotherapy. Secondly, the specific molecular mechanisms governing the relationship between CNKSR1 and TME in LUAD remain unknown. We would conduct further in-depth investigations into these aspects in our future research.

In conclusion, our study has identified CNKSR1 as a gene that strongly negatively correlated with PD-L1 expression and the infiltration of immune cells, particularly $CD8^+$ T cells in LUAD. By utilizing PD-L1 and CNKSR1 as biomarkers, we can demarcate distinct immune microenvironments. Unraveling the regulatory network underlying CNKSR1-related immune evasion may offer a promising pathway for the advancement of immunotherapeutic strategies for LUAD.

Ethics approval and consent to participate

This study was approved by the Ethics Committees of the Second Xiangya Hospital of Central South University. All enrolled participants gave written informed consent to participate and publish in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable.

Data availability statement

The datasets analyzed during the current study are publicly available. Further inquiries can be directed to the corresponding authors.

Funding

This study was supported by Natural Science Foundation of Changsha City (kq2202389, M.Peng), Natural Science Foundation of Hunan Provincial (2022JJ30831, M.Peng), the Scientific Research Launch Project for employees of the Second Xiangya Hospital (Q. Cai), and Health Research Project of Hunan Provincial Health Commission (W20243110, M.Peng).

CRediT authorship contribution statement

Qidong Cai: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Funding acquisition, Formal analysis, Data curation. **Mou Peng:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

All authors sincerely acknowledge the contributions from the TCGA project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e29126.

References

[1] A. Lahiri, A. Maji, P.D. Potdar, N. Singh, P. Parikh, B. Bisht, A. Mukherjee, M.K. Paul, Lung cancer immunotherapy: progress, pitfalls, and promises, Mol. Cancer 22 (2023) 1–37, https://doi.org/10.1186/s12943-023-01740-y.
[2] M. Reck, J. Remon, M.D. Hellmann, First-line immunotherapy for non–small-cell lung cancer, J. Clin. Oncol. 40 (2022) 586–597, https://doi.org/10.1200/JCO.21.01497.
[3] F. Zhou, M. Qiao, C. Zhou, The cutting-edge progress of immune-checkpoint blockade in lung cancer, Cell. Mol. Immunol. 18 (2021) 279–293, https://doi.org/10.1038/s41423-020-00577-5.
[4] A. Rittmeyer, F. Barlesi, D. Waterkamp, K. Park, F. Ciardiello, J. von Pawel, S.M. Gadgeel, T. Hida, D.M. Kowalski, M.C. Dols, D.L. Cortinovis, J. Leach, J. Polikoff, C. Barrios, F. Kabbinavar, O.A. Frontera, F. De Marinis, H. Turna, J.-S. Lee, M. Ballinger, M. Kowanetz, P. He, D.S. Chen, A. Sandler, D.R. Gandara, Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): a phase 3, open-label, multicentre randomised controlled trial, Lancet 389 (2017) 255–265, https://doi.org/10.1016/S0140-6736(16)32517-X.
[5] R.S. Herbst, P. Baas, D.W. Kim, E. Felip, J.L. Pérez-Gracia, J.Y. Han, J. Molina, J.H. Kim, C.D. Arvis, M.J. Ahn, M. Majem, M.J. Fidler, G. De Castro, M. Garrido, G.M. Lubiniecki, Y. Shentu, E. Im, M. Dolled-Filhart, E.B. Garon, Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial, Lancet 387 (2016) 1540–1550, https://doi.org/10.1016/S0140-6736(15)01281-7.
[6] H. Borghaei, L. Paz-Ares, L. Horn, D.R. Spigel, M. Steins, N.E. Ready, L.Q. Chow, E.E. Vokes, E. Felip, E. Holgado, F. Barlesi, M. Kohlhäufl, O. Arrieta, M. A. Burgio, J. Fayette, H. Lena, E. Poddubskaya, D.E. Gerber, S.N. Gettinger, C.M. Rudin, N. Rizvi, L. Crinò, G.R. Blumenschein, S.J. Antonia, C. Dorange, C. T. Harbison, F. Graf Finckenstein, J.R. Brahmer, Nivolumab versus docetaxel in advanced nonsquamous non–small-cell lung cancer, N. Engl. J. Med. 373 (2015) 1627–1639, https://doi.org/10.1056/nejmoa1507643.
[7] J. Brahmer, K.L. Reckamp, P. Baas, L. Crinò, W.E.E. Eberhardt, E. Poddubskaya, S. Antonia, A. Pluzanski, E.E. Vokes, E. Holgado, D. Waterhouse, N. Ready, J. Gainor, O. Arén Frontera, L. Havel, M. Steins, M.C. Garassino, J.G. Aerts, M. Domine, L. Paz-Ares, M. Reck, C. Baudelet, C.T. Harbison, B. Lestini, D.R. Spigel, Nivolumab versus docetaxel in advanced squamous-cell non–small-cell lung cancer, N. Engl. J. Med. 373 (2015) 123–135, https://doi.org/10.1056/nejmoa1504627.
[8] L. Fehrenbacher, A. Spira, M. Ballinger, M. Kowanetz, J. Vansteenkiste, J. Mazieres, K. Park, D. Smith, A. Artal-Cortes, C. Lewanski, F. Braiteh, D. Waterkamp, P. He, W. Zou, D.S. Chen, J. Yi, A. Sandler, A. Rittmeyer, Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial, Lancet (London, England) 387 (2016) 1837–1846, https://doi.org/10.1016/S0140-6736(16)00587-0.
[9] E.B. Garon, M.D. Hellmann, N.A. Rizvi, E. Carcereny, N.B. Leighl, M.J. Ahn, J.P. Eder, A.S. Balmanoukian, C. Aggarwal, L. Horn, A. Patnaik, M. Gubens, S. S. Ramalingam, E. Felip, J.W. Goldman, C. Scalzo, E. Jensen, D.A. Kush, R. Hui, Five-year overall survival for patients with advanced non-small-cell lung cancer treated with pembrolizumab: results from the phase i KEYNOTE-001 study, J. Clin. Oncol. 37 (2019) 2518–2527, https://doi.org/10.1200/JCO.19.00934.
[10] J. Wang, J. Sun, L.N. Liu, D.B. Flies, X. Nie, M. Toki, J. Zhang, C. Song, M. Zarr, X. Zhou, X. Han, K.A. Archer, T. O'Neill, R.S. Herbst, A.N. Boto, M.F. Sanmamed, S. Langermann, D.L. Rimm, L. Chen, Siglec-15 as an immune suppressor and potential target for normalization cancer immunotherapy, Nat. Med. 25 (2019) 656–666, https://doi.org/10.1038/s41591-019-0374-x.
[11] E.R. Parra, P. Villalobos, J. Zhang, C. Behrens, B. Mino, S. Swisher, B. Sepesi, A. Weissferdt, N. Kalhor, J.V. Heymach, C. Moran, J. Zhang, J. Lee, J. Rodriguez-Canales, D. Gibbons, I.I. Wistuba, Immunohistochemical and image analysis-based study shows that several immune checkpoints are Co-expressed in non-small cell lung carcinoma tumors, J. Thorac. Oncol. 13 (2018) 779–791, https://doi.org/10.1016/j.jtho.2018.03.002.
[12] J.L.F. Teh, A.E. Aplin, Arrested developments: CDK4/6 inhibitor resistance and alterations in the tumor immune microenvironment, Clin. Cancer Res. 25 (2019) 921–927, https://doi.org/10.1158/1078-0432.CCR-18-1967.
[13] Y.T. Liu, Z.J. Sun, Turning cold tumors into hot tumors by improving T-cell infiltration, Theranostics 11 (2021) 5265–5286, https://doi.org/10.7150/thno.58390.
[14] J. Galon, D. Bruni, Approaches to treat immune hot, altered and cold tumours with combination immunotherapies, Nat. Rev. Drug Discov. 18 (2019) 197–218, https://doi.org/10.1038/s41573-018-0007-y.
[15] D. Chen, G. Li, C. Ji, Q. Lu, Y. Qi, C. Tang, J. Xiong, J. Hu, F.B.A. Yasar, Y. Zhang, D.S.B. Hoon, Y. Yao, L. Zhou, Enhanced B7-H4 expression in gliomas with low PD-L1 expression identifies super-cold tumors, J. Immunother. Cancer 8 (2020) 1–10, https://doi.org/10.1136/jitc-2019-000154.
[16] S. Mariathasan, S.J. Turley, D. Nickles, A. Castiglioni, K. Yuen, Y. Wang, E.E.I.I.I. Kadel, H. Koeppen, J.L. Astarita, R. Cubas, S. Jhunjhunwala, R. Banchereau, Y. Yang, Y. Guan, C. Chalouni, J. Ziai, Y. Şenbabaoğlu, S. Santoro, D. Sheinson, J. Hung, J.M. Giltnane, A.A. Pierce, K. Mesh, S. Lianoglou, J. Riegler, R.A. D. Carano, P. Eriksson, M. Höglund, L. Somarriba, D.L. Halligan, M.S. van der Heijden, Y. Loriot, J.E. Rosenberg, L. Fong, I. Mellman, D.S. Chen, M. Green, C. Derleth, G.D. Fine, P.S. Hegde, R. Bourgon, T. Powles, TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells, Nature 554 (2018) 544–548, https://doi.org/10.1038/nature25501.

[17] P. Bischoff, A. Trinks, B. Obermayer, J.P. Pett, J. Wiederspahn, F. Uhlitz, X. Liang, A. Lehmann, P. Jurmeister, A. Elsner, T. Dziodzio, J.C. Rückert, J. Neudecker, C. Falk, D. Beule, C. Sers, M. Morkel, D. Horst, N. Blüthgen, F. Klauschen, Single-cell RNA sequencing reveals distinct tumor microenvironmental patterns in lung adenocarcinoma, Oncogene 40 (2021) 6748–6758, https://doi.org/10.1038/s41388-021-02054-3.
[18] K. Yoshihara, M. Shahmoradgoli, E. Martínez, R. Vegesna, H. Kim, W. Torres-Garcia, V. Treviño, H. Shen, P.W. Laird, D.A. Levine, S.L. Carter, G. Getz, K. Stemke-Hale, G.B. Mills, R.G.W. Verhaak, Inferring tumour purity and stromal and immune cell admixture from expression data, Nat. Commun. 4 (2013), https://doi.org/10.1038/ncomms3612.
[19] S.L. Carter, K. Cibulskis, E. Helman, A. McKenna, H. Shen, T. Zack, P.W. Laird, R.C. Onofrio, W. Winckler, B.A. Weir, R. Beroukhim, D. Pellman, D.A. Levine, E. S. Lander, M. Meyerson, G. Getz, Absolute quantification of somatic DNA alterations in human cancer, Nat. Biotechnol. 30 (2012) 413–421, https://doi.org/10.1038/nbt.2203.
[20] D. Aran, M. Sirota, A.J. Butte, Systematic pan-cancer analysis of tumour purity, Nat. Commun. 6 (2015) 1–11, https://doi.org/10.1038/ncomms9971.
[21] A.M. Newman, C.L. Liu, M.R. Green, A.J. Gentles, W. Feng, Y. Xu, C.D. Hoang, M. Diehn, A.A. Alizadeh, Robust enumeration of cell subsets from tissue expression profiles, Nat. Methods 12 (2015) 453–457, https://doi.org/10.1038/nmeth.3337.
[22] T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W.M. Mauck, Y. Hao, M. Stoeckius, P. Smibert, R. Satija, Comprehensive integration of single-cell data, Cell 177 (2019) 1888–1902.e21, https://doi.org/10.1016/j.cell.2019.05.031.
[23] R. Gao, S. Bai, Y.C. Henderson, Y. Lin, A. Schalck, Y. Yan, T. Kumar, M. Hu, E. Sei, A. Davis, F. Wang, S.F. Shaitelman, J.R. Wang, K. Chen, S. Moulder, S.Y. Lai, N.E. Navin, Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes, Nat. Biotechnol. 39 (2021) 599–608, https://doi.org/10.1038/s41587-020-00795-2.
[24] A.C. Habermann, A.J. Gutierrez, L.T. Bui, S.L. Yahn, N.I. Winters, C.L. Calvi, L. Peter, M.-I. Chung, C.J. Taylor, C. Jetter, L. Raju, J. Roberson, G. Ding, L. Wood, J.M.S. Sucre, B.W. Richmond, A.P. Serezani, W.J. McDonnell, S.B. Mallal, M.J. Bacchetta, J.E. Loyd, C.M. Shaver, L.B. Ware, R. Bremner, R. Walia, T. S. Blackwell, N.E. Banovich, J.A. Kropski, Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis, Sci. Adv. 6 (2020), https://doi.org/10.1126/sciadv.aba1972 eaba1972.
[25] C. Domínguez Conde, C. Xu, L.B. Jarvis, D.B. Rainbow, S.B. Wells, T. Gomes, S.K. Howlett, O. Suchanek, K. Polanski, H.W. King, L. Mamanova, N. Huang, P. A. Szabo, L. Richardson, L. Bolt, E.S. Fasouli, K.T. Mahbubani, M. Prete, L. Tuck, N. Richoz, Z.K. Tuong, L. Campos, H.S. Mousa, E.J. Needham, S. Pritchard, T. Li, R. Elmentaite, J. Park, E. Rahmani, D. Chen, D.K. Menon, O.A. Bayraktar, L.K. James, K.B. Meyer, N. Yosef, M.R. Clatworthy, P.A. Sims, D.L. Farber, K. Saeb-Parsy, J.L. Jones, S.A. Teichmann, Cross-tissue immune cell analysis reveals tissue-specific features in humans, Science 376 (2022) eabl5197, https://doi.org/10.1126/science.abl5197.
[26] A.B. Jaffe, A. Hall, A. Schmidt, Association of CNK1 with Rho guanine nucleotide exchange factors controls signaling specificity downstream of Rho, Curr. Biol. 15 (2005) 405–412, https://doi.org/10.1016/j.cub.2004.12.082.
[27] A.M. Dahlin, M.L. Henriksson, B. Van Guelpen, R. Stenling, Å. Öberg, J. Rutegrd, R. Palmqvist, Colorectal cancer prognosis depends on T-cell infiltration and molecular characteristics of the tumor, Mod. Pathol. 24 (2011) 671–682, https://doi.org/10.1038/modpathol.2010.234.
[28] P. Jiang, S. Gu, D. Pan, J. Fu, A. Sahu, X. Hu, Z. Li, N. Traugh, X. Bu, B. Li, J. Liu, G.J. Freeman, M.A. Brown, K.W. Wucherpfennig, X.S. Liu, Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response, Nat. Med. 24 (2018) 1550–1558, https://doi.org/10.1038/s41591-018-0136-1.
[29] J. Yuan, T. Cai, X. Zheng, Y. Ren, J. Qi, X. Lu, H. Chen, H. Lin, Z. Chen, M. Liu, S. He, Q. Chen, S. Feng, Y. Wu, Z. Zhang, Y. Ding, W. Yang, Potentiating CD8(+) T cell antitumor activity by inhibiting PCSK9 to promote LDLR-mediated TCR recycling and signaling, Protein Cell 12 (2021) 240–260, https://doi.org/10.1007/s13238-021-00821-2.
[30] M. Desbois, A.R. Udyavar, L. Ryner, C. Kozlowski, Y. Guan, M. Dürrbaum, S. Lu, J.P. Fortin, H. Koeppen, J. Ziai, C.W. Chang, S. Keerthivasan, M. Plante, R. Bourgon, C. Bais, P. Hegde, A. Daemen, S. Turley, Y. Wang, Integrated digital pathology and transcriptome analysis identifies molecular mediators of T-cell exclusion in ovarian cancer, Nat. Commun. 11 (2020), https://doi.org/10.1038/s41467-020-19408-2.
[31] K. Nishiyama, M. Maekawa, T. Nakagita, J. Nakayama, T. Kiyoi, M. Chosei, A. Murakami, Y. Kamei, H. Takeda, Y. Takada, S. Higashiyama, CNKSR1 serves as a scaffold to activate an EGFR phosphatase via exclusive interaction with RhoB-GTP, Life Sci. Alliance 4 (2021) 1–22, https://doi.org/10.26508/LSA.202101095.
[32] A. Fischer, T. Brummer, B. Warscheid, G. Radziwill, Differential tyrosine phosphorylation controls the function of CNK1 as a molecular switch in signal transduction, Biochim. Biophys. Acta Mol. Cell Res. 1853 (2015) 2847–2855, https://doi.org/10.1016/j.bbamcr.2015.08.014.
[33] C.U. Blank, J.B. Haanen, A. Ribas, T.N. S, The "cancer immunogram,", Science 352 (2016) 658–660.
[34] J.M. Taube, R.A. Anders, G.D. Young, H. Xu, R. Sharma, T.L. McMiller, S. Chen, A.P. Klein, D.M. Pardoll, S.L. Topalian, L. Chen, Colocalization of inflammatory response with B7-H1 expression in human melanocytic lesions supports an adaptive resistance mechanism of immune escape, Sci. Transl. Med. 4 (2012), https://doi.org/10.1126/scitranslmed.3003689.
[35] M.J. Marinissen, J.S. Gutkind, Scaffold proteins dictate Rho GTPase-signaling specificity, Trends Biochem. Sci. 30 (2005) 423–426, https://doi.org/10.1016/j.tibs.2005.06.006.
[36] A. Clapéron, M. Therrien, KSR and CNK: two scaffolds regulating RAS-mediated RAF activation, Oncogene 26 (2007) 3143–3158, https://doi.org/10.1038/sj.onc.1210408.
[37] R.D. Fritz, G. Radziwill, CNK1 and other scaffolds for Akt/FoxO signaling, Biochim. Biophys. Acta Mol. Cell Res. 1813 (2011) 1971–1977, https://doi.org/10.1016/j.bbamcr.2011.02.008.
[38] J. Gil, A. Ramírez-Torres, S. Encarnación-Guevara, Lysine acetylation and cancer: a proteomics perspective, J. Proteonomics 150 (2017) 297–309, https://doi.org/10.1016/j.jprot.2016.10.003.
[39] M.C. Mendoza, E.E. Er, J. Blenis, The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation, Trends Biochem. Sci. 36 (2011) 320–328, https://doi.org/10.1016/j.tibs.2011.03.006.
[40] A. Fischer, W.W.D. Mühlhäuser, B. Warscheid, G. Radziwill, Membrane localization of acetylated CNK1 mediates a positive feedback on RAF/ERK signaling, Sci. Adv. 3 (2017), https://doi.org/10.1126/sciadv.1700475.
[41] R.D. Fritz, G. Radziwill, CNK1 promotes invasion of cancer cells through NF-kappaB-dependent signaling, Mol. Cancer Res. 8 (2010) 395–406, https://doi.org/10.1158/1541-7786.MCR-09-0296.
[42] A.B. Jaffe, P. Aspenström, A. Hall, Human CNK1 acts as a scaffold protein, linking Rho and Ras signal transduction pathways, Mol. Cell Biol. 24 (2004) 1736–1746, https://doi.org/10.1128/mcb.24.4.1736-1746.2004.
[43] D. Li, A.M. Miermont, R. Sable, H.S. Quadri, L.A. Mathews Griner, S.E. Martin, T. Odzorig, S. De, M. Ferrer, A.S. Powers, S.M. Hewitt, U. Rudloff, Scaffolding protein connector enhancer of kinase suppressor of Ras 1 (CNKSR1) regulates MAPK inhibition responsiveness in pancreas cancer via crosstalk with AKT signaling, Mol. Cancer Res. 21 (2023) 316–331, https://doi.org/10.1158/1541-7786.MCR-21-1036.
[44] M. Indarte, R. Puentes, M. Maruggi, N.T. Ihle, G. Grandjean, M. Scott, Z. Ahmed, E.J. Meuillet, S. Zang, R. Lemos, L. Du-Cuny, F.I.A.L. Layng, R.G. Correa, L. A. Bankston, R.C. Liddington, L. Kirkpatrick, G. Powis, An inhibitor of the pleckstrin homology domain of CNK1 selectively blocks the growth of mutant KRAS cells and tumors, Cancer Res. 79 (2019) 3100–3111, https://doi.org/10.1158/0008-5472.CAN-18-2372.
[45] Y. Zhang, Y. Lin, D. Lv, X. Wu, W. Li, X. Wang, D. Jiang, Identification and validation of a novel signature for prediction the prognosis and immunotherapy benefit in bladder cancer, PeerJ 10 (2022) 1–27, https://doi.org/10.7717/peerj.12843.
[46] J. Chen, H. Jin, H. Zhou, X. Hei, K. Liu, Research into the characteristic molecules significantly affecting liver cancer immunotherapy, Front. Immunol. 14 (2023) 1–13, https://doi.org/10.3389/fimmu.2023.1029427.
[47] Y. Chen, X. Zhou, Y. Xie, J. Wu, T. Li, T. Yu, Y. Pang, W. Du, Establishment of a seven-gene signature associated with CD8(+) T cells through the utilization of both single-cell and bulk RNA-sequencing techniques in clear cell renal cell carcinoma, Int. J. Mol. Sci. 24 (2023), https://doi.org/10.3390/ijms241813729.
[48] S. Leng, G. Nie, C. Yi, Y. Xu, L. Zhang, L. Zhu, Machine learning - derived identification of tumor - infiltrating immune cell - related signature for improving prognosis and immunotherapy responses in patients with skin cutaneous melanoma, Cancer Cell Int. (2023) 1–15, https://doi.org/10.1186/s12935-023-03048-9.

[49] M. Yi, D.V. Nissley, F. McCormick, R.M. Stephens, ssGSEA score-based Ras dependency indexes derived from gene expression data reveal potential Ras addiction mechanisms with possible clinical implications, Sci. Rep. 10 (2020) 1–16, https://doi.org/10.1038/s41598-020-66986-8.
[50] J. Novaković, P. Strbac, D. Bulatović, Toward optimal feature selection using ranking methods and classification algorithms, Yugosl. J. Oper. Res. 21 (2011) 119–135, https://doi.org/10.2298/YJOR1101119N.
[51] R.J. Urbanowicz, R.S. Olson, P. Schmitt, M. Meeker, J.H. Moore, Benchmarking relief-based feature selection methods for bioinformatics data mining, J. Biomed. Inf. 85 (2018) 168–188, https://doi.org/10.1016/j.jbi.2018.07.015.